UNIVERSITY OF CALCUTTA

IS THE WEAK LAW OF LARGE NUMBERS REALLY WEAK?

Submitted in partial fulfillment of the requirements for the award of the degree of
BACHELOR OF SCIENCE (HONS.) IN STATISTICS

Submitted by:
Name: SAYANTAN PAL
CU Roll No.: 223146-21-0083
CU Reg. No.: 146-1111-0504-22
Semester: VI
Paper: DSE-B2

Under the guidance of
Dr. Tuhinsubhra Bhattacharya
Assistant Professor
Department of Statistics
MAULANA AZAD COLLEGE

<u>Certificate</u>

This is to certify that the project paper titled **"IS THE WEAK LAW OF LARGE NUMBERS REALLY WEAK?"** has been submitted by ***Sayantan Pal*** in partial fulfillment of the requirements for the B.Sc. (Hons) degree in Statistics from University of Calcutta.

The research work presented in this project has been conducted under my guidance and supervision. I can confirm that the results and findings presented in this project are the outcome of investigator's efforts and the research work conducted during the specified period.

Furthermore, to the best of my knowledge, the results and findings reported in this project have not been submitted for the award of any other degree or diploma in any academic institution.

I recommend this project paper for evaluation and fulfillment of the academic requirements for the B.Sc. (Hons) in Statistics degree. If you have any further queries or require additional information, please do not hesitate to contact me.

Dated: ------------------------------------

(Tuhinsubhra Bhattacharya)
Department of Statistics
Maulana Azad College

Declaration

I, ***Sayantan Pal*** , a student of Semester VI, entitled to the programme B.Sc.(Hons) in Statistics at ***Maulana Azad College***, affiliated to University of Calcutta with ***Registration No. 146-1111-0504-22*** and ***Roll No. 223146-21-0083***, solemnly declare that the project titled ***"IS THE WEAK LAW OF LARGE NUMBERS REALLY WEAK?***" is a genuine and original work undertaken by me under the supervision of Prof. Tuhinsubhra Bhattacharya as a part of my Bachelor's programme.

I further affirm that:

• The analysis, interpretations and conclusions presented in this project are the results of my own research and statistical analysis.

• The project has not been submitted to any other academic degree or assessment at any institution.

• I understand the academic integrity requirements of my university, and I have adhered to all the guidelines and regulations regarding plagiarism and research ethics.

• In case of any unintentional oversight error, I take full responsibility for the same and will co-operate with the necessary revisions or corrections as advised by my faculty or supervisor.

• I am aware of the consequences of the academic misconduct, and I affirm that this project is entirely free of any act of dishonesty or deception.

I am sincerely committed to the successful completion of this project.

(Sayantan Pal)

Abstract

This project investigate the validity of ***Khinchin's weak law of large numbers (WLLN)*** beyond the classical framework of independent and identically distributed (i.i.d.) sequence of random variables. *In this project, we investigate : what happens if the sequence of random variables are just independent (but not identical)*? we find that the weak law of large numbers (WLLN) still holds in such cases, as long as the sequence of random variables have finite means and the overall variability is in control- that ensure convergence in probability of the sample average to a limiting mean.

Through theoretical analysis and illustrative examples, this projects demonstrate that independence, combined with certain regularity conditions (e.g, Lindeberg type or Vanishing average variance) is sufficient for WLLN to hold . Here we also use simulation to investigate the rate of convergence of the sample mean to the population mean and we will draw some further conclusion.

This relaxation makes the law more flexible & useful in real world situations where conditions like identical distribution don't always apply.

Acknowledgement

I would like to express my heartfelt gratitude and extend my sincerest appreciation to all the individuals who have played a major role in the successful completion of this project. Their guidance, support and expertise have been invaluable throughout my journey. Firstly, I am deeply indebted to my supervisor ***Prof. Tuhinsubhra Bhattacharya*** for his unwavering dedication, insightful feedback and continuous encouragement.

I would like to extend my sincere thanks to ***Prof. Partha Pal***, Head of the Statistics department and ***Prof. Anup Kumar Giri*** for their constant support and encouragement. I am very grateful for their valuable insights which has played a pivotal role in shaping my project.

I would like to acknowledge my fellow classmates who have supported me throughout this journey with their small but valuable advices. I would like to convey heartfelt appreciation to my seniors, ***Subhendu Ghosh & Tathastu Ghosh*** , for offering me substantial support throughout this journey.

Lastly, I would like to thank Maulana Azad College for providing me a nurturing academic environment and the necessary resources to pursue my project.

Contents

Introduction

"Absence of evidence is not evidence of absence.": – *Carl Sagan*.

The **Weak Law of Large Numbers (WLLN)** stands as a cornerstone of probability theory and statistical inference, providing a rigorous foundation for the principle that empirical averages tend to reflect theoretical expectations as the sample size grows. In its classical form, due to **A.Ya. Khinchin**, the WLLN asserts that the sample mean of a sequence of **independent and identically distributed (i.i.d.) random variables** with a finite expectation converges **in probability** to the population mean. This convergence justifies the use of sample averages as consistent estimators in statistical practice and underpins the reliability of empirical methods across a wide range of disciplines.

However, the assumptions of **identical distribution and independence**, while mathematically convenient, often fail to hold in practical settings—such as in econometrics, signal processing, and data streams—where observations may be **independent but not identically distributed (i.n.i.d.)**. In such cases, the classical WLLN no longer applies directly, and convergence of the sample mean is not guaranteed without additional structural conditions.

This project undertakes a rigorous study of the Weak Law of Large Numbers beyond the i.i.d. framework. The primary objective is to explore the extent to which the WLLN remains valid under relaxed distributional assumptions. We begin by reviewing Khinchin's WLLN and its foundational proof, then systematically analyze scenarios where the i.i.d. condition is weakened to independence with non-identical distributions. Key results, including generalizations under bounded variance, Counter examples are also presented to highlight the necessity of such conditions.

By bridging classical theory with more general stochastic frameworks, this study aims to provide a comprehensive understanding of the convergence behavior of sample means under varying degrees of distributional heterogeneity—an issue of both theoretical significance and practical relevance in modern probabilistic modeling.

Objectives

The primary aim of this project is to explore and critically analyze the Weak Law of Large Numbers (WLLN), focusing on both its classical formulation and its extensions to more general settings. The specific objectives are as follows:

1. **To present a rigorous exposition of Khinchin's Weak Law of Large Numbers**, including a detailed proof under the standard assumptions of independence and identical distribution, with finite mean.

2. **To investigate the theoretical limitations of the classical WLLN** by identifying the essential role of the i.i.d. assumption and exploring cases where this condition is violated.

3. **To examine the behavior of the sample mean under sequences of independent but non-identically distributed (i.n.i.d.) random variables**, identifying conditions under which convergence in probability still holds.

4. **To analyze various generalizations of the WLLN**, including:
   - The Lindeberg and Lyapunov conditions.
   - Kolmogorov's version of the WLLN for independent sequences with uniformly bounded variance.

5. **To construct and analyze counter-examples** that demonstrate the failure of the WLLN when key assumptions (such as boundedness of variance or asymptotic regularity) are not satisfied.

6. **To provide illustrative examples and simulations**, where appropriate, to support theoretical results and enhance understanding of the convergence behavior of sample means in different stochastic environments.

7. **To contextualize the findings within practical applications**, such as statistical estimation, econometric modeling, and data science, where the assumption of identically distributed data often does not hold.

Basic Terminologies

Few basic terminologies in this context are given below –

- ***Different modes of convergence:***

  Consider a sequence $\{X_n\}, n \geq 1$ of RV's defined on a probability space (Ω, A, P).

  The different modes of convergence are:

  a) convergence in probability

  b) convergence in distribution or law

  c) convergence in r^{th} mean

  d) Almost sure convergence

a. convergence in probability:

A sequence of random variables $\{X_n\}, n \geq 1$ is said to be converges in probability to a random variable X if, for every ε >0

$$P(|X_n - X| > \varepsilon) \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty$$

Notation: $X_n \xrightarrow{P} X$

b. Convergence in law:

A sequence of random variables $\{X_n\}, n \geq 1$ is said to be **converges in distribution (in law)** to a random variable X, written as:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

where $F_{X_n}(x)$ is the CDF of X_n.

- ***A sufficient condition of convergence In probability:***

  Let $\{X_n\}, n \geq 1$ be a sequence of random variables such that

  E$(X_n) \rightarrow \theta$ and ***var***$(X_n) \rightarrow 0$ as $n \rightarrow \infty$

  Then, $X_n \xrightarrow{P} \theta$ in probability as $n \rightarrow \infty$.

Weak Law Of Large Numbers(WLLN):

Let $\{X_n\}, n \geq 1$ be a sequence of random variable and **S_n=$\sum_k X_k$=nth partial sum**.

We say that $\{X_n\}$ obeys weak law of large numbers wrt a sequence of real no.$\{b_n\}$, $b_n > 0$ & b_n increases as n increases , $b_n \to \infty$ as $n \to \infty$ .If there exist another sequence of real no. $\{a_n\}$ such that

$\frac{S_n - a_n}{bn} \xrightarrow{p} 0$ as $n \to \infty$,

Where **a_n** is called **centering constant** & **b_n** is called **norming constant.**

- ***Khinchin's WLLN:***

  Let $\{X_n\}$ be a sequence of i.i.d. random variables with **finite expected value** μ=E$[X_n]$. Define the sample average:

$$\bar{X} = \frac{1}{n} \sum_k X_k$$

  Then , $\bar{X} \xrightarrow{P}$ μ.

  That is, the sample average **converges in probability** to the expected value μ as $n \to \infty$.

Theoretical Analysis

The classical ***Weak Law of Large Numbers (WLLN)****, attributed to* ***A.Y. Khinchin****, asserts that the sample average of independent and identically distributed (i.i.d.) random variables with finite mean converges in probability to the common mean. However, the assumption of identical distribution can be stringent in practical applications. This study aims to determine under what relaxed conditions—specifically without requiring identical distributions—the WLLN continues to hold.*

- ***Theoretical Framework:***

  Let $\{X_n\}, n \geq 1$ be a sequence of independent random variable (but not necessarily identically distributed) with finite means μ_i=E(X_i) and variances $\sigma_i{}^2 = var(X_i)$.

  Define the sample average:

$$\bar{X} = \frac{1}{n}\sum_i X_i$$

  We are interested in the convergence in probability:

$$\bar{X} \xrightarrow{P} \mu \text{ as } n \to \infty \text{, for some constant } \mu \in \mathbb{R}.$$

> ***Theorem:***
>
> Let $\{X_n\}, n \geq 1$ be a sequence of independent random variables such that:
>
> 1. $\sup_i (var(X_i)) \leq C < \infty$
> 2. $\frac{1}{n}\sum_i E(X_i) \to \mu \in \mathbb{R}$.
>
> Then, $\bar{X} \xrightarrow{P} \mu$.

[*Intuition behind the bounded variance condition:*

Uniformly bounded variances prevent any single term from dominating the average due to large fluctuations. Without this control, a few variables with large variance could distort the sample average even as the number of terms increases. This boundedness ensures the sample average stabilizes around the expected mean behavior.]

Proof:

Let , $\mu_n = \frac{1}{n}\sum_i E(X_i)$.Then,

$P(|\bar{X}_n - \mu| > \varepsilon) = P(|(\bar{X}_n - \mu_n) + (\mu_n - \mu)| > \varepsilon)$

$\leq P(|\bar{X}_n - \mu_n| > \frac{\varepsilon}{2}) + P(|\mu_n - \mu| > \frac{\varepsilon}{2})$ (By Triangle Law)

Since, $\mu_n \to \mu$ as $n \to \infty$, $P(|\mu_n - \mu| > \frac{\varepsilon}{2}) = 0$

Again, $P(|\bar{X}_n - \mu_n| > \frac{\varepsilon}{2}) \leq \frac{4}{\varepsilon^2} Var(\bar{X}_n)$ (By Chebyshev's Inequality)

$\leq \frac{4c}{\varepsilon^2 n} \to 0$ as $n \to \infty$

Thus, $P(|\bar{X}_n - \mu_n| > \frac{\varepsilon}{2}) = 0$

Therefore, $P(|\bar{X}_n - \mu| > \varepsilon) = 0$

Which implies , $\bar{X}_n \xrightarrow{P} \mu$ QED

Explanation Using Some Basic Theoretical Examples

Consider the following examples,

Example.1: (*Decreasing variance and constant means*)

Set up: Let $\{X_n\}, n \geq 1$ be a sequence of independent **Normal** random variables such that

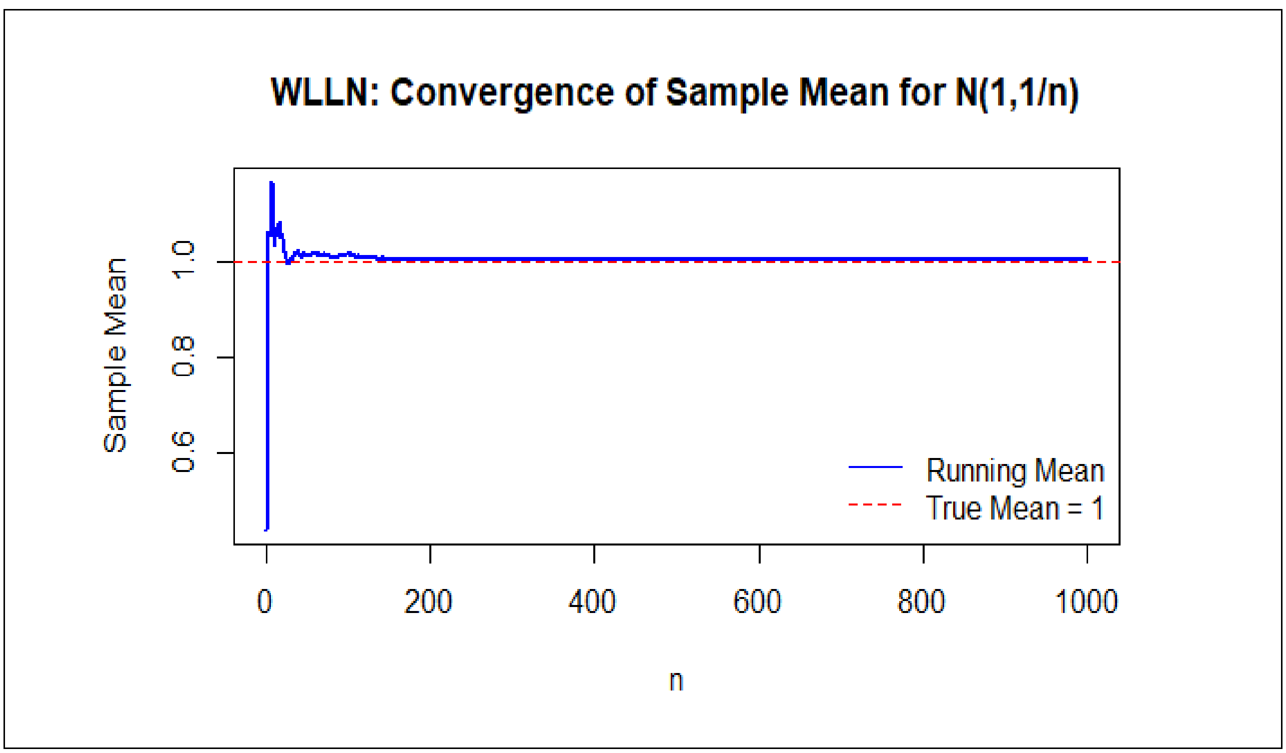
E(X_i) = 1 and ***Var***(X_i) = $\frac{1}{i}$

Clearly , ***E***($\bar{X}_n$) = 1

$$\textbf{\textit{Var}}(\bar{X}_n) = \frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) = \frac{1}{n^2}\sum_i \frac{1}{i} = \ln(n)/n^2 \rightarrow 0 \quad \text{as} \quad n\rightarrow\infty$$

Then by sufficient condition of convergence :

$$\bar{X}_n \xrightarrow{P} 1$$



Example.2: (*Linearly increasing variance and constant means*)

Set up: Let $\{X_n\}, n \geq 1$ be a sequence of independent **Normal** random variables such that

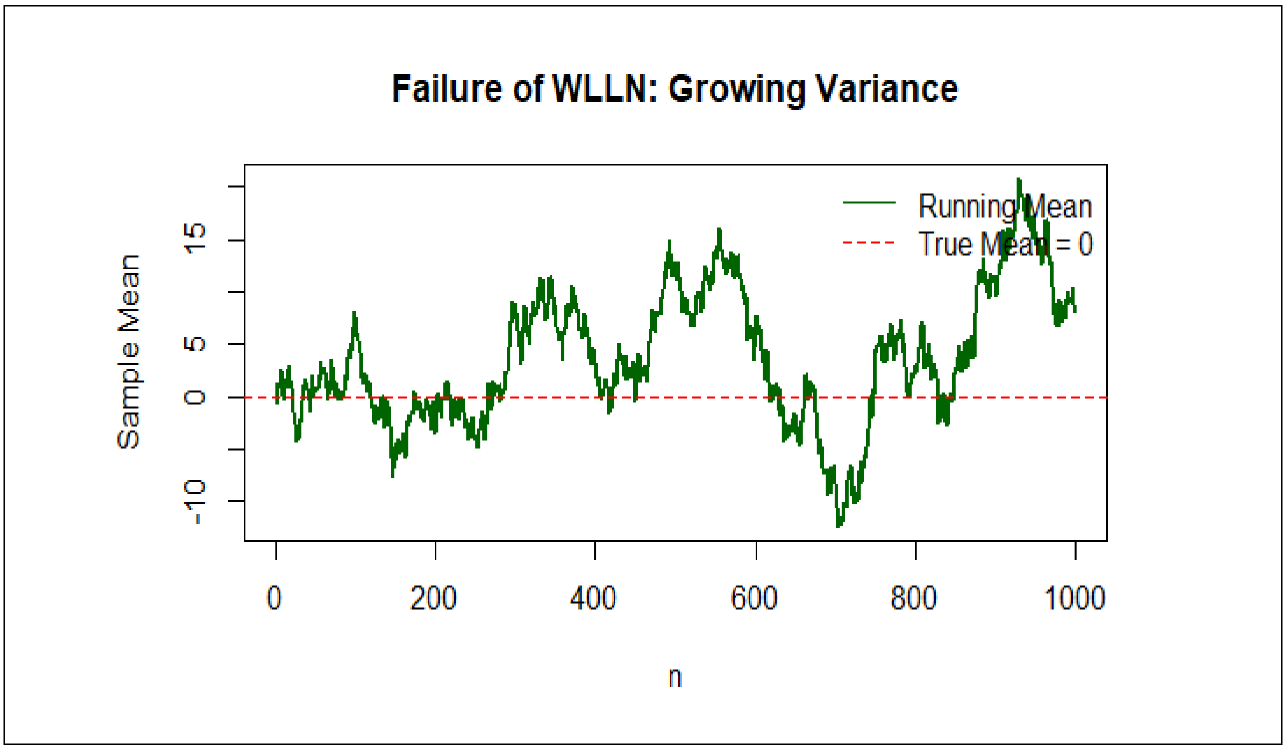
E(X_i) = 0 and ***Var***(X_i) = i

Clearly , ***E***($\bar{X}_n$) = 0

$$\boldsymbol{Var}(\bar{X}_n) = \frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) = \frac{1}{n^2}\sum_i i = \frac{n+1}{2} \rightarrow \infty \quad \text{as } n \rightarrow \infty$$

Then we can say that ,

$$\bar{X}_n \overset{P}{\nrightarrow} 0$$



<u>Example.3</u>: (*<u>Slowly increasing variance and constant means</u>*)

<u>Set up:</u> Let $\{X_n\}, n \geq 1$ be a sequence of independent **Normal** random variables such that

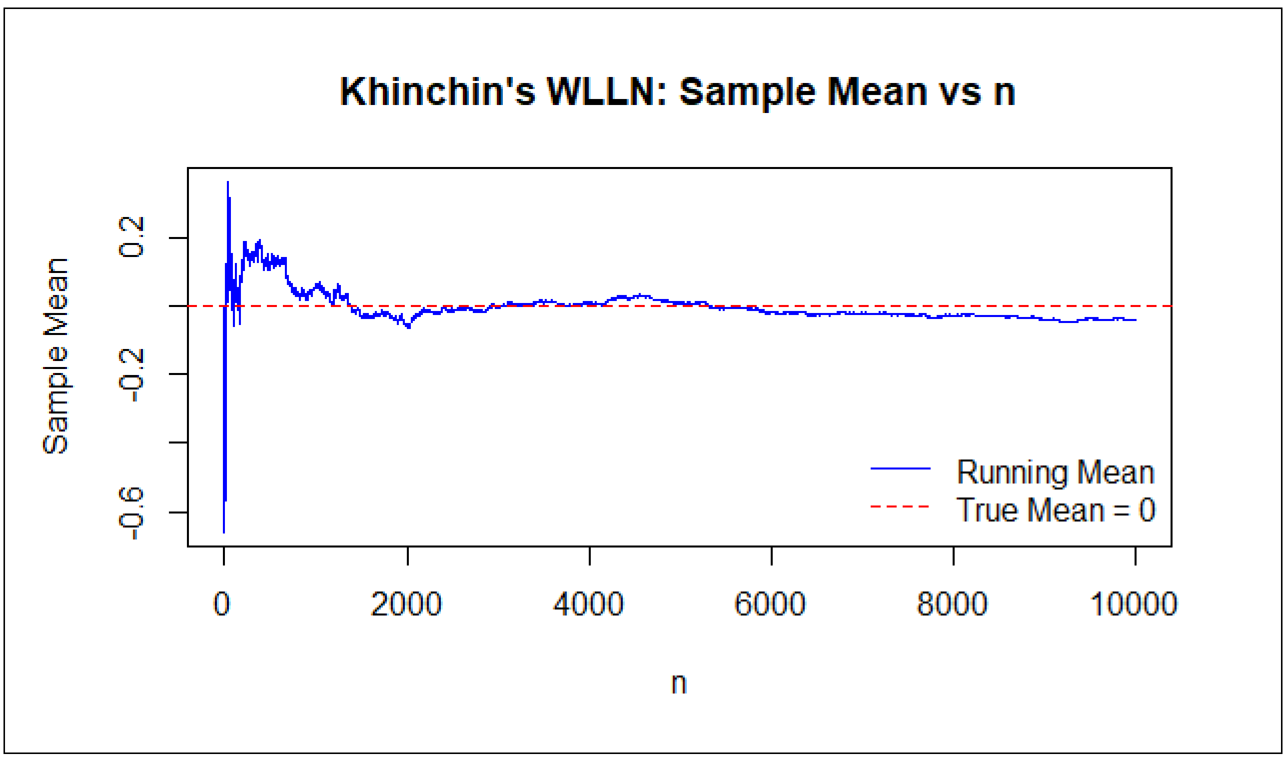
E(X_i) = 0 and ***Var***(X_i) = $ln(i)$

Clearly , ***E***($\bar{X}_n$) = 0

$$\boldsymbol{Var}(\bar{X}_n) = \frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) = \frac{1}{n^2}\sum_i ln(i) = \frac{1}{n^2}\int_1^n ln(x)dx \sim \frac{ln(n)}{n} \rightarrow 0 \quad \text{as } n\rightarrow\infty$$

Then by sufficient condition of convergence :

$$\bar{X}_n \xrightarrow{P} 0$$



Example.4: (*Alternating means and constant variance*)

Set up: Let $\{X_n\}, n \geq 1$ be a sequence of independent **Normal** random variables such that

E(X_i) = $\frac{(-\mathbf{1})^{\boldsymbol{i}}}{\boldsymbol{i}}$ and ***Var***(X_i) = 2

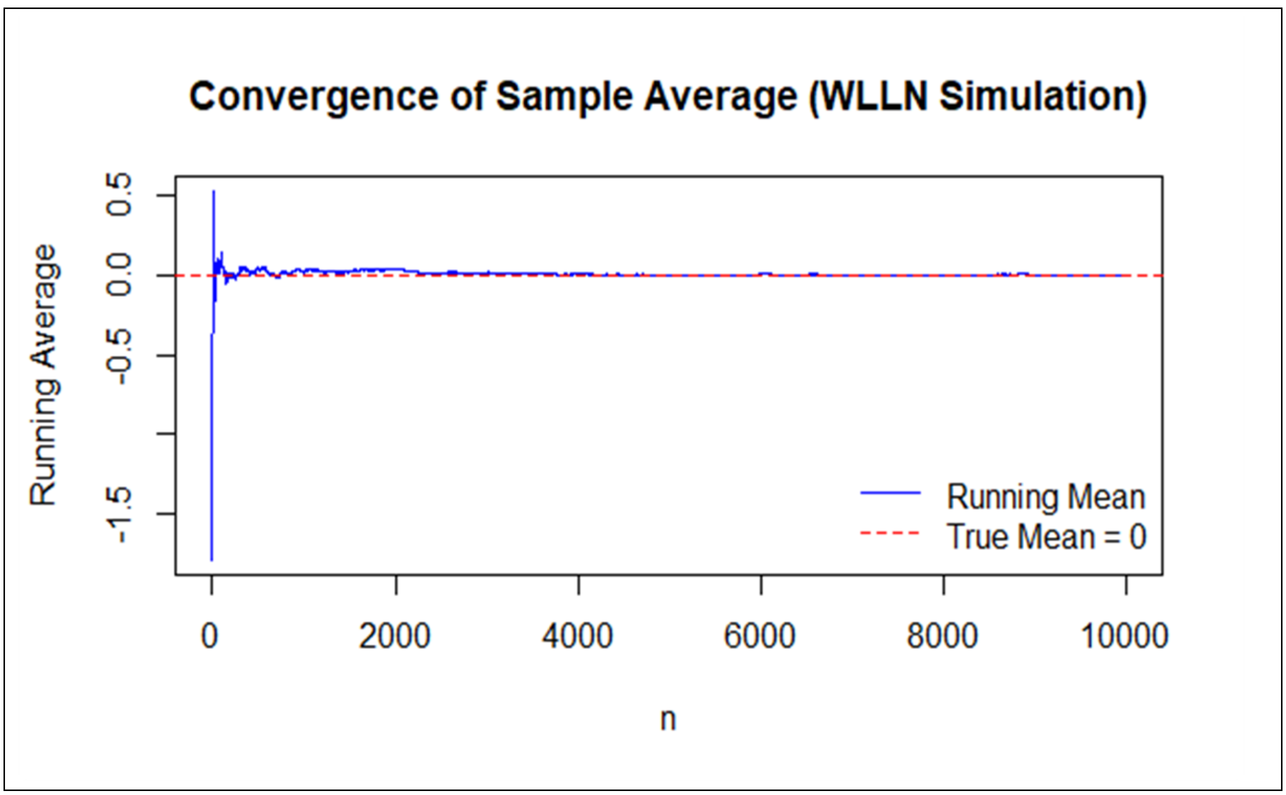
Clearly , ***E***($\bar{X}_n$) = $\frac{1}{n}\sum_i \boldsymbol{E}(Xi) = \frac{1}{n}\sum_i \frac{(-\mathbf{1})^{\boldsymbol{i}}}{\boldsymbol{i}} = \frac{ln(2)}{n} \to 0$ as $n \to \infty$

(from the expansion of ln(1+x) series)

$$\boldsymbol{Var}(\bar{X}_n) = \frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) = \frac{1}{n^2}\sum_i 2 = \frac{2n}{n^2} \to 0 \quad \text{as } n \to \infty$$

Then by sufficient condition of convergence:

$$\bar{X}_n \xrightarrow{P} 0$$



Example.5: (*Uniformly bounded support*)

Set up: Let $\{X_n\}, n \geq 1$ be a sequence of independent **Uniform**(-1+$\frac{1}{n}$, $1-\frac{1}{n}$) random variables.

Consider $a_n = -1 + \frac{1}{n}$ and $b_n = 1 - \frac{1}{n}$ then,

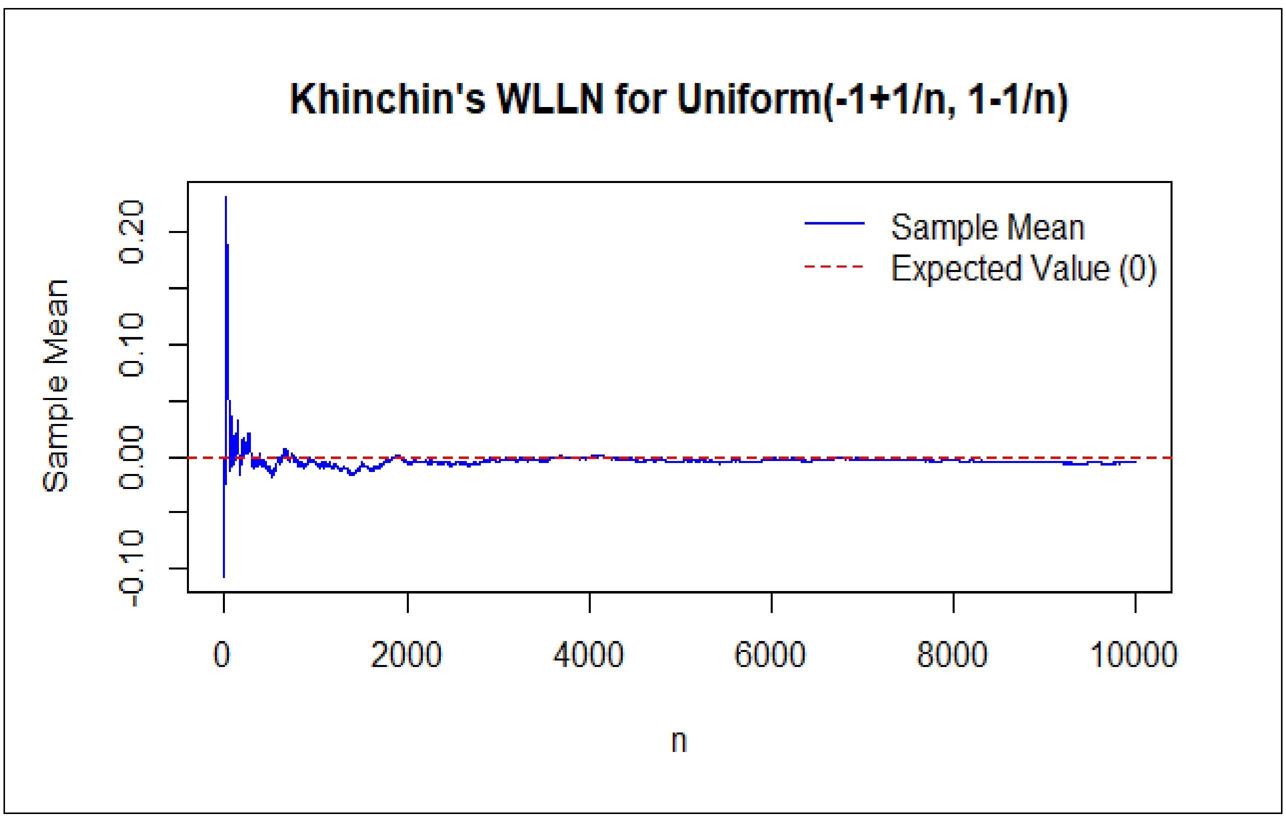
$\boldsymbol{E}(X_i) = \frac{a_n+b_n}{2} = 0$ and $\boldsymbol{Var}(X_i) = \frac{(b_n-a_n)^2}{12} = \frac{(2-\frac{2}{n})^2}{12} = \frac{4(1-\frac{1}{n})^2}{12} \to \frac{1}{3}$ as $n \to \infty$

Clearly , $\boldsymbol{E}(\bar{X}_n) = \frac{1}{n}\sum_i \boldsymbol{E}(Xi) = \frac{1}{n}\sum_i 0 \to 0$ as $n \to \infty$

$$\boldsymbol{Var}(\bar{X}_n) = \frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) \leq \frac{1}{n^2}\sum_i \frac{1}{3} = \frac{n}{3n^2} \to 0 \quad \text{as } n \to \infty$$

Then by sufficient condition of convergence :

$$\bar{X}_n \xrightarrow{P} 0$$



Example.6: *(Increasing variance and increasing mean)*

Set up: Let $\{X_n\}, n \geq 1$ be a sequence of independent **Exponential** random variables with parameter $\lambda_n = (1 + \frac{1}{n})$; then,

E(X_i) = $\frac{1}{\lambda_i}$ and ***Var***(X_i) = $\frac{1}{{\lambda_i}^2}$

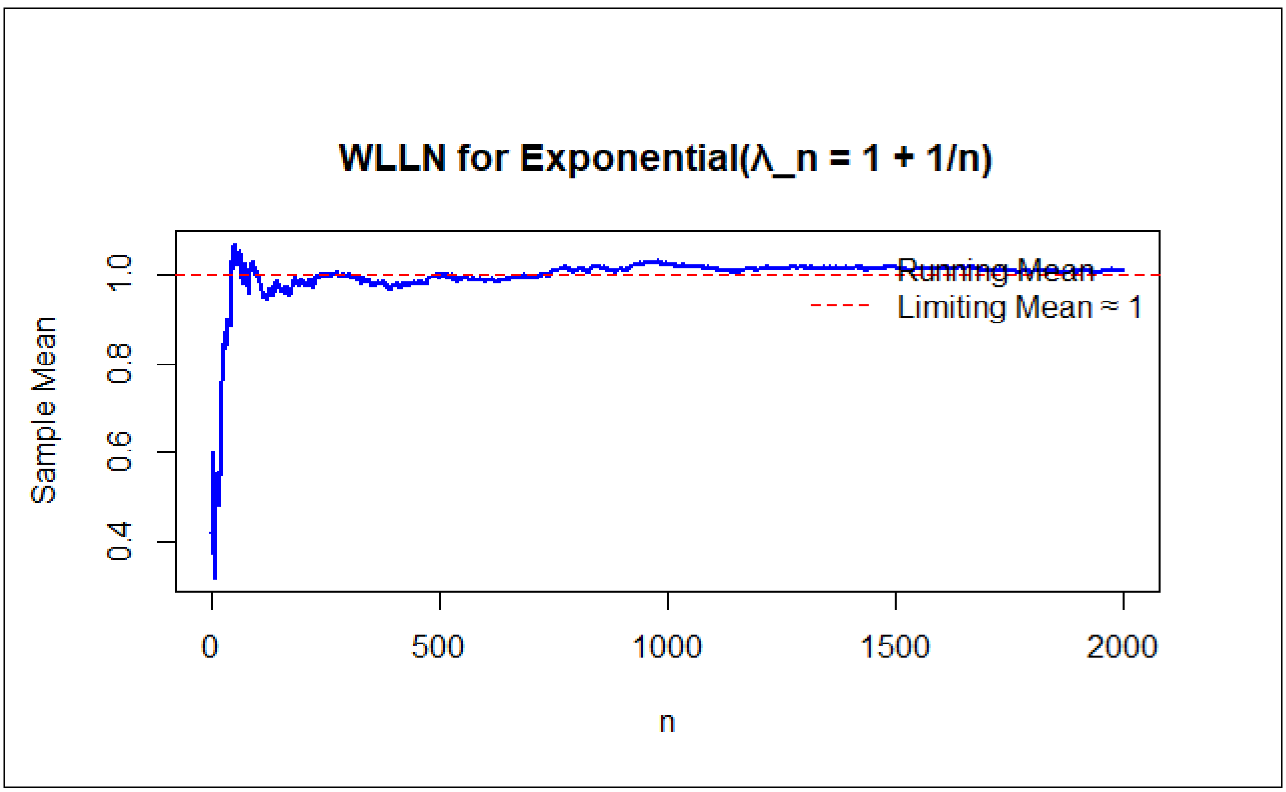
Clearly , ***E***($\bar{X}_n$) = $\frac{1}{n}\sum_i \boldsymbol{E}(Xi)$ = $\frac{1}{n}\sum_i \frac{1}{\lambda_i}$ = $\frac{1}{n}\sum_i \frac{i}{1+i}$ $= 1 - \frac{1}{n}\sum_i \frac{1}{1+i}$

= 1- $\frac{1}{n}\int_1^n \frac{1}{1+x}dx \approx 1 - \frac{ln(n+1)}{n} \to 1$ as n→∞

Var($\bar{X}_n$) = $\frac{1}{n^2}\sum_i \boldsymbol{var}(Xi)$ = $\frac{1}{n^2}\sum_i (\frac{i}{1+i})^2 \leq \frac{1}{n^2}\sum_i 1 = \frac{1}{n} \to 0$ as n→∞

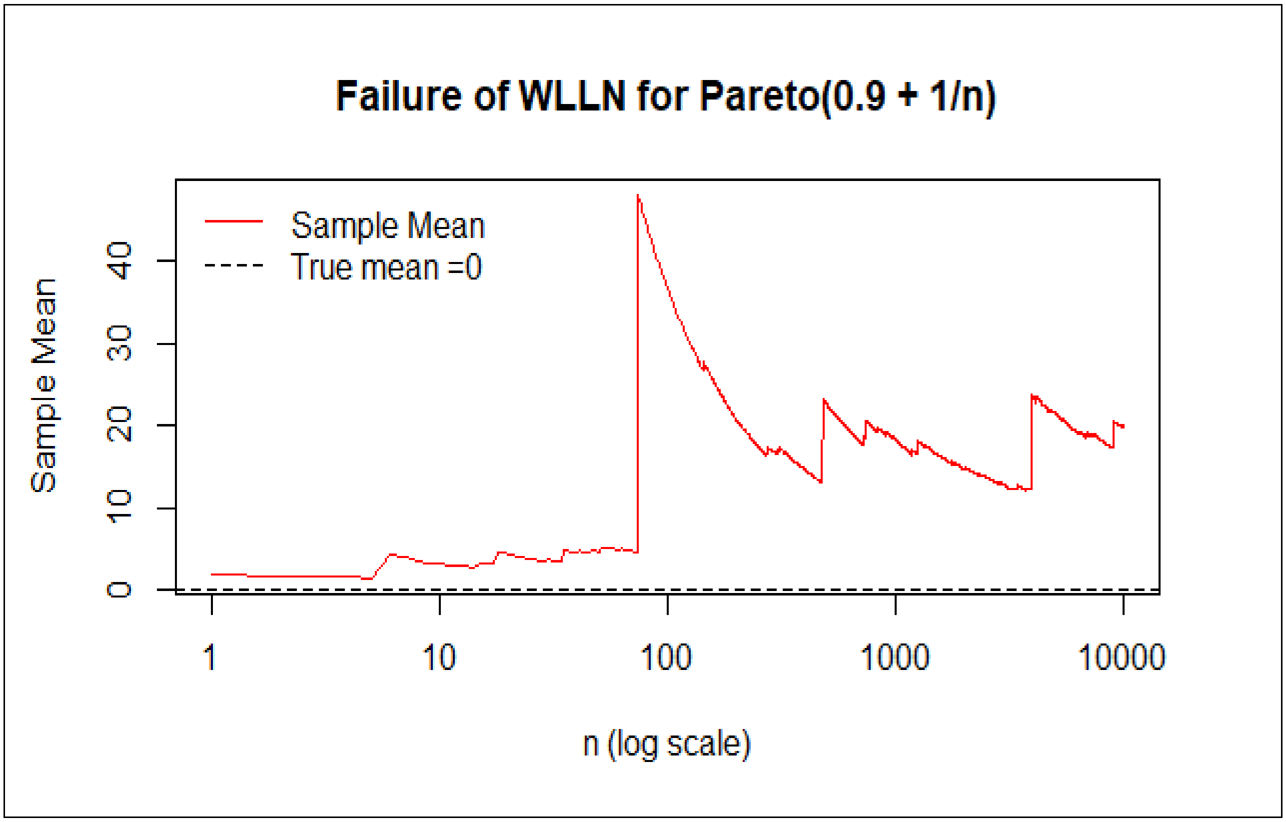
Then by sufficient condition of convergence :

$$\bar{X}_n \xrightarrow{P} 1$$



<u>Example.7</u>:(*<u>Increasing variance and increasing mean</u>*)

<u>Set up:</u> Let $\{X_n\}, n \geq 1$ be a sequence of independent **Pareto** random variables with parameter α_n *where* $\alpha_n = 0.9 + \frac{1}{n}$. Note that, for a pareto distribution with shape parameter $\alpha_n > 0$,expectation of X_n exists if $\alpha_n > 1$. But here $\alpha_n \rightarrow 0.9$ as n→∞. Therefore mean of this distribution doesn't exists. Hence, khinchin's WLLN doesn't hold in this set up due to the violation on the condition **finite expectation** (Since, the expectation is unbounded here).



Example.8:(*Increasing shape and increasing scale*)

Set up: Let $\{X_n\}, n \geq 1$ be a sequence of independent **Gamma** random variables with parameter $\alpha_n = n \; and \; p_n = n$; then,

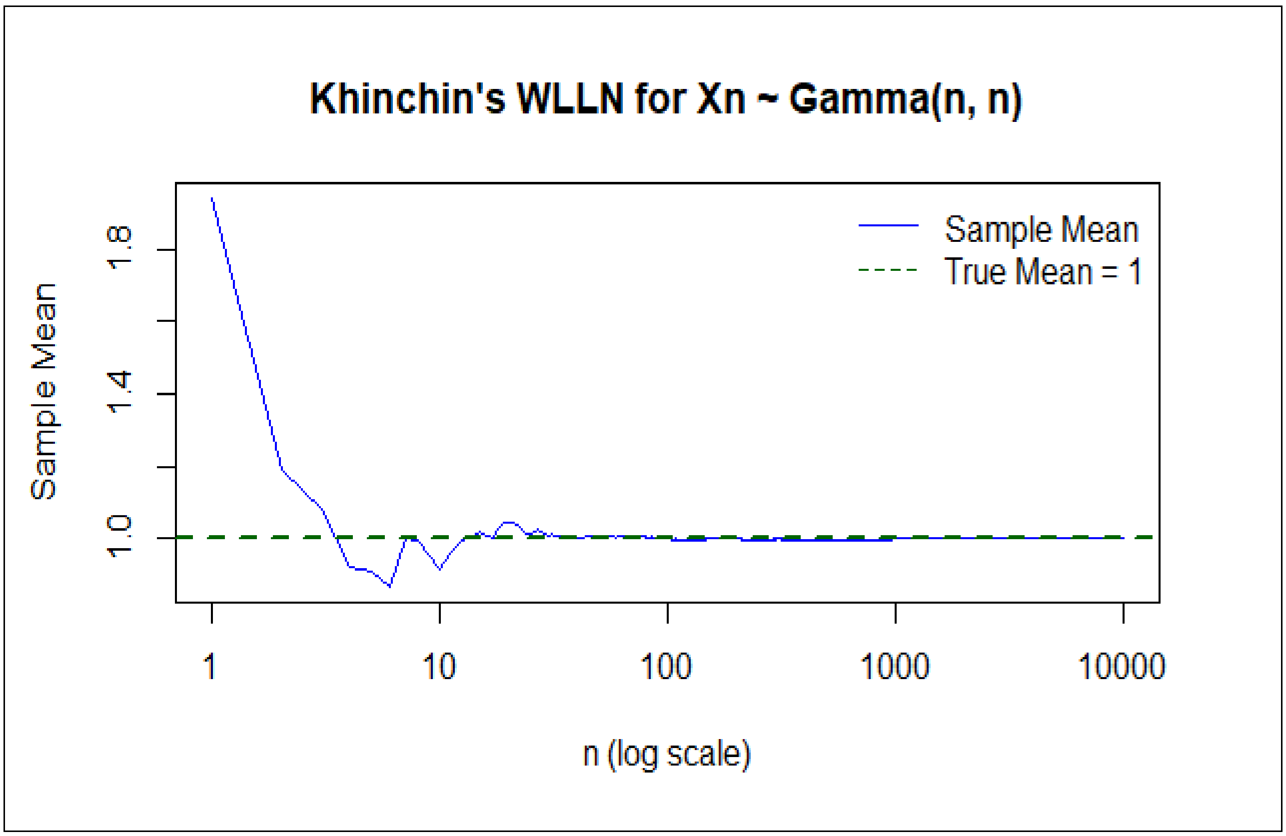
E(X_i) = $\frac{p_n}{\alpha_n} = 1$ and ***Var***(X_i) = $\frac{p_n}{{\alpha_n}^2} = \frac{1}{n}$

Clearly , ***E***($\bar{X}_n$) = $\frac{1}{n}\sum_i \boldsymbol{E}(Xi) = \frac{1}{n}\sum_i 1 \rightarrow 1$ as n→∞

Var($\bar{X}_n$)= $\frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) = \frac{1}{n^2}\sum_i \frac{1}{i} \approx \frac{1}{n^2}\int_1^n \frac{1}{x} dx = \frac{ln(n)}{n^2} \rightarrow 0$ as n→∞

Then by sufficient condition of convergence:

$$\bar{X}_n \xrightarrow{P} 1$$



<u>Example.9</u>:(<u>*Increasing variance and constant mean*</u>)

<u>Set up:</u> Let $\{X_n\}, n \geq 1$ be a sequence of independent **Lognormal** random variables with parameter $\mu_n = 0 \; and \; \sigma_n{}^2 = 1 - \frac{1}{n}$; then,

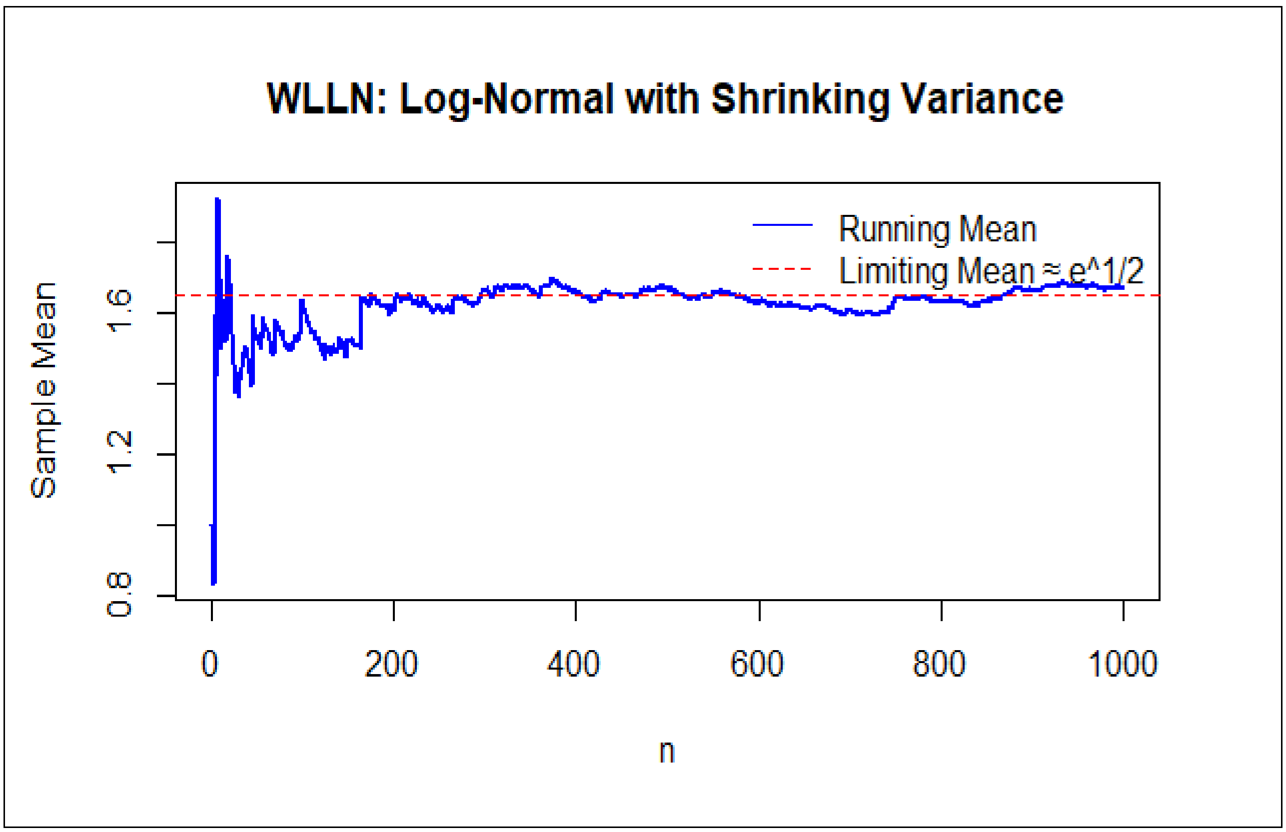
E(X_i) = $e^{\mu_n + \frac{\sigma_n{}^2}{2}} = e^{0+\frac{n-1}{2n}} = e^{\frac{1}{2}-\frac{1}{n}}$ and ***Var***(X_i) = $(e^{\sigma_n{}^2} - 1)(\, e^{2\mu_n + \sigma_n{}^2}) = (e^{1-\frac{1}{n}} - 1)(\, e^{1-\frac{1}{n}})$

Clearly , ***E***($\bar{X}_n$) = $\frac{1}{n}\sum_i \boldsymbol{E}(Xi) = \frac{1}{n}\sum_i e^{\frac{1}{2}-\frac{1}{i}} \to e^{\frac{1}{2}}$ as n→∞

Var($\bar{X}_n$) = $\frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) = \frac{1}{n^2}\sum_i (e^{1-\frac{1}{i}} - 1)(\, e^{1-\frac{1}{i}}) \to 0$ as n→∞

Then by sufficient condition of convergence :

$$\bar{X}_n \xrightarrow{P} e^{\frac{1}{2}}$$



Example.10: (*Scaled Chi-square variable*)

Set up: Let $\{X_n\}, n \geq 1$ be a sequence of independent random such that

$X_n = \frac{1}{\sqrt{n}} Z_n$ where $Z_n \sim \chi^2{}_1$ $distribution$

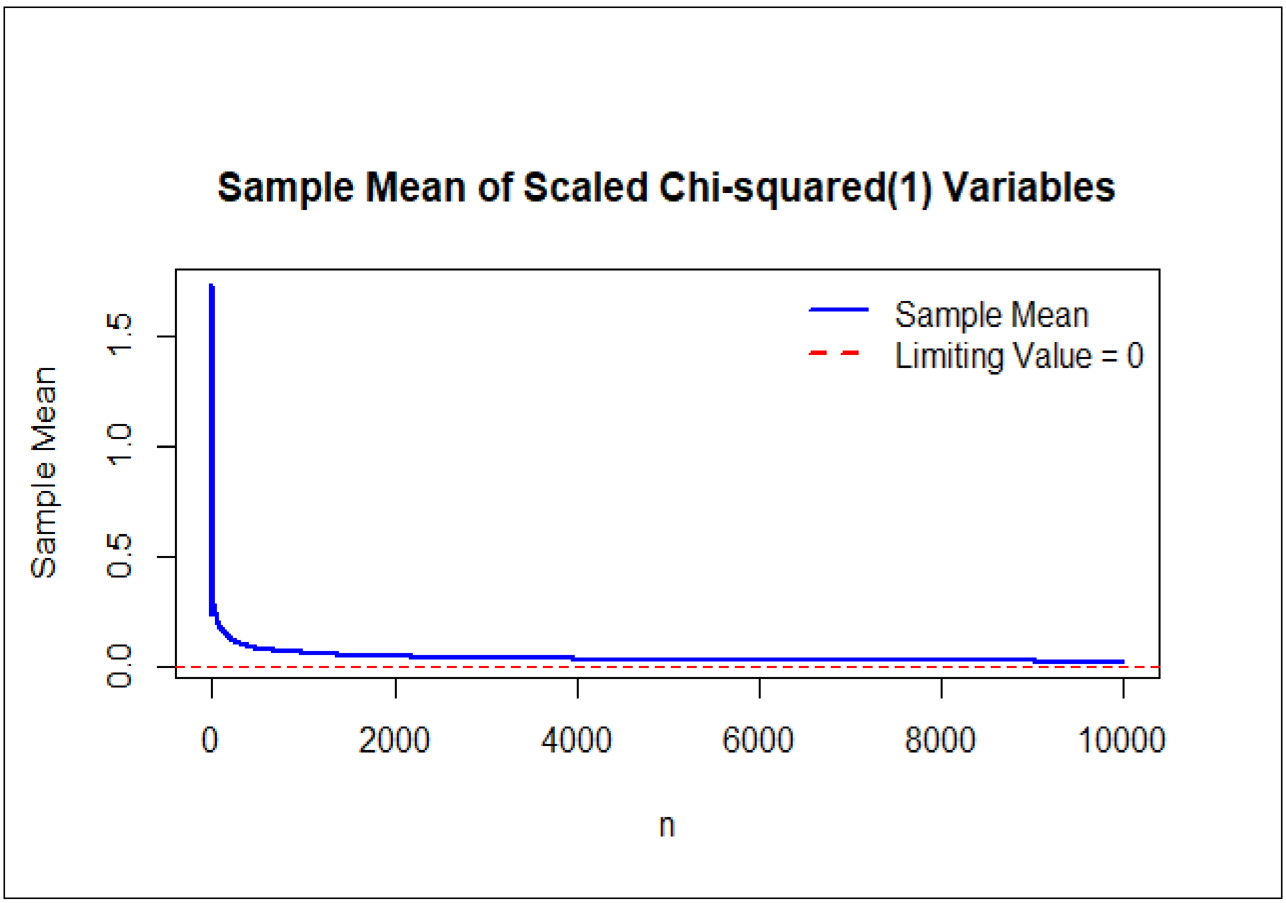
Then $\boldsymbol{E}(X_i) = \frac{1}{\sqrt{i}}$ and $\boldsymbol{Var}(X_i) = \frac{2}{i}$

Clearly , $\boldsymbol{E}(\bar{X}_n) = \frac{1}{n}\sum_i \boldsymbol{E}(Xi) = \frac{1}{n}\sum_i \frac{1}{\sqrt{i}} \approx \frac{1}{n}\int_1^n \frac{1}{\sqrt{x}} dx = \frac{2}{\sqrt{n}} \rightarrow 0$ as $n \rightarrow \infty$

$$\boldsymbol{Var}(\bar{X}_n) = \frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) = \frac{1}{n^2}\sum_i \frac{2}{i} \leq \frac{2ln(n)}{n^2} \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty$$

Then by sufficient condition of convergence:

$$\bar{X}_n \xrightarrow{P} 0$$



Observations

No.	Distribution Setup	Mean ($E(\bar{X}_n)$)	Variance ($Var(\bar{X}_n)$)	WLLN Holds?	Reason
1	$N\left(1, \frac{1}{n}\right)$	1	$\frac{ln(n)}{n^2}$	YES	Variance tends to 0 too fast
2	$N(0, n)$	0	$\frac{n+1}{2}$	NO	Variance grows to infinity
3	$N(1, ln(n))$	1	$\frac{ln(n)}{n}$	YES	Variance is bounded and convergent
4	$N\left(\frac{(-1)^n}{n}, 2\right)$	ln(2)	$\frac{2}{n}$	YES	Bounded Variance
5	R$(-1+\frac{1}{n}, 1-\frac{1}{n})$	0	$\frac{1}{3n}$	YES	Bounded Support
6	Exp(1+$\frac{1}{n}$)	$1-\frac{ln(n+1)}{n}$	$\frac{1}{n}$	YES	Converges to Exp(1)
7	Pareto (0.9+$\frac{1}{n}$)	∞(For early terms)	Infinte for small n	NO	Heavy tails, infinite mean
8	Gamma(n,n)	1	$\frac{ln(n)}{n^2}$	YES	Variance is bounded
9	Λ(0,1-$\frac{1}{n}$)	$e^{\frac{1}{2}}$	Zero for large n	YES	Converges to Λ(0,1)
10	Scaled ${\chi^2}_1$	$\frac{2}{\sqrt{n}}$	Zero for large n	YES	Finite Mean and bounded Variance

Findings

The objective of this analysis was to verify **Khinchin's Weak Law of Large Numbers (WLLN)** across a variety of sequences of independent, but non-identically distributed random variables. Each sequence was defined by a distinct distributional form and parameterization that affected either the mean, the variance, or both.

The central theme of WLLN is that the sample average

$$\bar{X} = \frac{1}{n}\sum_k X_k$$

should converge in probability to a constant (usually the mean), even when the random variables are not identically distributed, provided certain conditions are met, particularly relating to the behavior of their variances.

Summary of Observations:

1. **Sequences with Decreasing Variance**:
   Examples like $\boldsymbol{N}$ (1 , $\frac{1}{n}$), **Gamma** (n , n) and scaled chi-square sequences $X_n = \frac{1}{\sqrt{n}} Z_n$ where $Z_n \sim \boldsymbol{\chi^2}_{\mathbf{1}}$ $distribution$, all showed rapid decay in variance. In these cases, the sample mean concentrated tightly around the expected value. This aligns with WLLN, where convergence is promoted by diminishing variance.
2. **Sequences with Constant but Finite Variance**:
   Uniform and Lognormal distributions (e.g., **Uniform**(-1+$\frac{1}{n}$, $1-\frac{1}{n}$) , **Λ**(0,1-$\frac{1}{n}$)) showed bounded variances. The mean values were stable across the sequence, and the Law held as variance didn't increase with n .
3. **Slowly Increasing Variance**:
   In the case of $\boldsymbol{N}$ (1 , ln(n)) although variance increased with n, it did so logarithmically, which is slow enough that the average variance $\frac{1}{n^2}\sum_i ln(i)$ = $\frac{1}{n^2}\int_1^n ln(x)dx$ ~ $\frac{ln(n)}{n}$ → 0 .

This was sufficient for WLLN to hold — a powerful illustration that identical distribution is not necessary, and even mild non-uniformity in variance can be tolerated.

4. **Sequences with Linearly or Super-Linearly Increasing Variance**: In the case of $\boldsymbol{N}(0,n)$, variance increased linearly with n, resulting in a sample mean with variance $\sim \frac{1}{n}$, which does not converge to zero fast enough. Here, the sample mean showed high variability and instability, and WLLN failed to hold.
5. **Heavy-Tailed Distributions**: The Pareto sequence with shape parameter $\alpha_n = 0.9 + \frac{1}{n}$ had infinite mean or infinite variance in initial terms. Despite convergence of the parameter, early terms dominate and disrupt convergence. Hence, WLLN fails, highlighting the sensitivity to tail behavior and integrability conditions.
6. **Converging Means but Stable Variance**: The sequence $\boldsymbol{N}\left(\frac{(-1)^n}{n}, 2\right)$, had alternating means converging to 0 and fixed variance. Here, the sample mean was stable, and WLLN held — emphasizing that mean convergence + bounded variance is sufficient, even when individual X_n vary in sign or shift.
7. **Exponential Family Distributions:** The exponential sequence with parameter $\lambda_n = (1 + \frac{1}{n})$ approached standard exponential behavior. Variance and mean both converged, leading to consistent behavior of the sample mean and validating WLLN.

Key Theoretical Takeaways

- **Identically Distributed is Not Required**:
  Khinchin's WLLN does not require identical distributions — the independence and the asymptotic behavior of the variance are the keys.
- **Sufficient Conditions:**
  The Lindeberg-type conditions and Lyapunov criteria provide more general conditions under which WLLN can hold for non-iid sequences. One practical test is:

  If $\frac{1}{n^2}\sum_i \boldsymbol{var}(Xi) \to 0$ as $n \to \infty$, then $\bar{X}_n \xrightarrow{P} \mu$.

- **Failure Cases Reveal Sensitivity:**
  Examples with infinite variance or heavy tails (e.g., Pareto) demonstrate that even a few "bad" variables in the sequence can prevent convergence. This underlines the importance of uniform integrability and finite second moments.

Inferential Aspects Based on the Central Limit Theorem in the i.n.i.d. Setup:

While **Khinchin's WLLN** ensures that the sample mean converges in probability to the population mean, but it says nothing about the **distributional shape** of the sample mean. This is where the **Central Limit Theorem (CLT)** plays a key role: it explains **how** the sample mean behaves in distribution, allowing for **normal approximation**—a cornerstone of statistical inference.

Lindeberg–Feller CLT (Generalized CLT)

For independent, non-identically distributed sequence of random variables $\{X_n\}, n \geq 1$ with:

- $E(X_i) = \mu_i$
- $var(X_i) = {\sigma_i}^2$
- *Finite second moment ,i.e.,* $E(X_i^2) < \infty$

  define:

- $S_n = \sum_k (X_k - \mu_k)$
- ${v_n}^2 = \sum_k {\sigma_k}^2$

  Then, under **Lindeberg's condition** (which controls the influence of large individual terms), we have:

$$\frac{S_n}{V_n} \xrightarrow{d} N(0,1)$$

Inferential Implications:

1. **Confidence Intervals**:
   The CLT justifies constructing **approximate normal confidence intervals** for

the sum or average of i.n.i.d. variables, provided the total variance is finite and Lindeberg's condition holds. That is:

$$\bar{X}_n \sim N\left(\frac{1}{n}\sum_i \mu_i, \frac{1}{n^2}\sum_i \sigma_i^2\right)$$

enabling inference even in heterogeneous settings.

2. **Hypothesis Testing**:
   In classical tests (e.g. one-sample z-test), the CLT allows for standardization of test statistics when sampling from i.n.i.d. populations, provided total variance is well-behaved.

3. **Robustness to Heterogeneity**:
   The CLT accommodates varying distributions, unlike simpler versions which require identical distributions. This supports inference in fields like econometrics or medical studies, where variability in measurement conditions is common.

4. **Finite-Sample Adjustments**:
   While asymptotic normality holds under mild conditions, practitioners must assess **variance dominance**—i.e., whether any single term contributes disproportionately to total variance. If so, normal approximation may be poor, requiring techniques like the bootstrap.

5. **Comparison with WLLN**:
   - **WLLN (Khinchin)** ensures consistency of sample mean.
   - **CLT** enables **approximate distributional inference** — confidence intervals and hypothesis tests.

Together, they form the foundation of **frequentist inference**: WLLN ensures estimators converge; CLT explains their distribution for large samples.

Future Scope of the Project

The future scope of this project are as follows:

- ***Violation of Independent Condition Instead of Identical condition:***

  In this project, we verified the **Khinchin's Weak Law of Large Numbers (WLLN)** under a variety of **independent but not identically distributed** settings. A natural and theoretically compelling extension is to investigate the behavior of sample averages when the assumption of **independence** is relaxed, but the sequence remains **identically distributed**. That is, we consider random variables $X_1, X_2,\ldots$ which are identically distributed, but not necessarily independent. This relaxation gives rise to a variety of dependent structures commonly encountered in real-world applications, such as **time series data, spatial data, network data,** and **econometrics.**

  Key Directions for Future Exploration

1. **Weak Dependence Structures:**
   - Investigate WLLN under mixing conditions (α-mixing, ϕ-mixing, ρ-mixing) or strong mixing sequences.
   - For example, when dependence between X_i and X_{i+k} weakens as k→ ∞, WLLN may still hold.
   - Study conditions like: $\sum_k \alpha(k)^{\frac{\delta}{2+\delta}} < \infty$ for some $\delta > 0$, where α(k) denotes the mixing coefficient.
2. **Martingale Differences:**
   - Explore cases where X_n forms a **martingale difference sequence** (MDS), which models many practical phenomena.
   - WLLN holds for MDS under certain boundedness or integrability conditions.
   - This is especially useful in econometrics and online learning.
3. **Markov Chains:**
   - Examine the sample mean behavior when the X_n form a Markov chain.
   - Under ergodicity and stationarity, many forms of the law of large numbers still hold.
   - Particularly relevant to Monte Carlo methods (MCMC) and stochastic simulations.

4. **Exchangeable Sequences:**
    - Study WLLN under **exchangeability**, where joint distributions are invariant under permutations.
    - Although not independent, de Finetti's theorem provides a representation linking them to i.i.d. sequences.
5. **Functional Dependence (Non-linear Time Series)**:
    - Explore settings such as:

$$X_n = g(\varepsilon_n, \varepsilon_{n-1}, \ldots)$$

      where ε_n are i.i.d., and g defines a dependence structure.
    - Examples include ARCH/GARCH models used in financial volatility modeling.

Theoretical and Practical Significance:

- **Broadens Applicability:**
  Real-world data often exhibit temporal or spatial dependence (e.g., stock prices, weather data). Understanding WLLN in these contexts makes the theory applicable beyond classical settings.
- **Links with Ergodic Theory:** The study leads into **ergodic theorems**, a generalization of the WLLN for stationary stochastic processes.
- **Strengthens Statistical Inference:** Ensuring that averages still converge under dependence validates many estimators and test statistics used in modern statistics.
- **Potential for Simulations:** The simulation framework developed in this project can be extended to test WLLN under dependent conditions using **auto-correlated or Markovian structures**.
  The extension to not **independent but identically distributed sequences** provides a natural next step that brings Khinchin's WLLN closer to practical modeling scenarios. This direction bridges classical probability with modern stochastic processes and opens rich opportunities for both **theoretical advancement and real-world data analysis.**

➢ *Transformation-Based Approach to Restore i.i.d.-Like Behavior:*

While many real-world random sequences are **not i.i.d.**, in some cases, a transformation (e.g., normalization, standardization, rescaling) can recover an approximate **i.i.d. behavior** from a non-identically distributed or even weakly dependent sequence. This is particularly relevant when the goal is to apply laws like **WLLN** or the **Central Limit Theorem (CLT)**, which are well-understood in the i.i.d. setting.

The central idea is:
Even if a sequence is not i.i.d., can we find a transformation or normalization under which the sample average behaves similarly to that of an i.i.d. sequence?

This direction hinges critically on the **rate of convergence of means** and **variances** of the underlying random variables.

Theoretical Framework:

Assume $X_1, X_2, \ldots\ldots$ is a sequence of independent (but not identically distributed) random variables with means μ_n and variances $\sigma_n{}^2$ .Then define:

- Centered version: $Y_n = X_n - \mu_n$
- Normalized version: $Z_n = \frac{X_n - \mu_n}{\sigma_n}$
- If the sequence Z_n has uniformly bounded moments, and the average variance behaves well (e.g., $\frac{1}{n^2}\sum_i \sigma_i{}^2$) then the sample mean may still converge **as if** it came from an i.i.d. structure.

Future Research Directions

1.**Explore transformation classes**:

- Study transformations $T_n(X_n)$ such that the transformed variables converge to i.i.d. or ergodic behavior.
- Useful for applying standard tools like CLT or strong law to more general setups.

2.Establish necessary convergence rates:

- Quantify how fast the mean $\mu_n \to \mu$ and variance $\sigma_n{}^2 \to \sigma^2$ must converge to mimic i.i.d. behavior.
- Investigate bounds on $|\mu_n - \mu|$ and variance $|\sigma_n{}^2 - \sigma^2|$ that ensure convergence of $\bar{X}_n \xrightarrow{P} \mu$.

3. **Uniform Integrability and Tail Control:**
- Determine when the transformations lead to uniformly integrable sequences.
- Focus on controlling the tails to ensure transformation doesn't introduce extreme variance.

4. **Simulation Validation:**
- Extend your simulation code to apply transformations (e.g., z-scores or empirical normalization) to sequences and test convergence.
- Compare transformed sample means to i.i.d. benchmarks to validate behavior.

5. **Applications in Statistical Estimation:**
- Estimators such as sample mean, regression coefficients, and bootstrap estimators may retain consistency when input sequences are normalized appropriately.
- Can lead to robust inference methods in heteroscedastic or non-homogeneous data environments.

Conceptual Contribution

This insight contributes a powerful lens:
- Instead of insisting on **i.i.d. assumptions**, we can **engineer i.i.d.-like behavior** through careful preprocessing or understanding of asymptotics.
- It shifts focus from the structure of the data to **the rate of convergence of its properties**, which is more general and applicable to complex datasets.

Final Perspective

The future scope of this project can therefore extend into **transformation theory, empirical process convergence,** and **generalized laws of large numbers** for structured but non-i.i.d. data. It offers a bridge between **asymptotic probability theory** and **realistic statistical modeling**, opening doors to research in:
- Functional analysis of random sequences
- Data-dependent transformations in statistics
- Asymptotics in machine learning and non-parametric statistics

Conclusion

The experiment illustrates **The power and limits of Khinchin's WLLN** in the context of independence without identical distribution. It demonstrates that convergence of the sample mean is intimately tied to how variances behave collectively. Even sequences with divergent individual variances may yield convergence if the aggregate variance of the sample mean vanishes.

This nuanced understanding is vital in modern applications, such as:

- Online learning and streaming data, where data points come from different distributions.
- Monte Carlo simulations, which often rely on weak law behavior under non-i.i.d. sampling.
- Heavy-tailed models in finance, where convergence properties can mislead if assumptions are not properly verified.

  This analysis therefore emphasizes that variance control, not distributional uniformity, is the true cornerstone of convergence in the WLLN framework.

Bibliography

The books which were used to develop theories in this project are:-

[1] George Casella and Roger Berger. Statistical Inference. Duxbury Thomson Learning, 2002.

[2] Joseph Blitzstein and Jessica Hwang. Introduction to Probability. CHAPMAN & HALL/CRC Press ,2019.

[3] A Mood F Graybill and D Boes. Introduction to the theory of statistics. McGraw Hill, 1974.

[4] Robert Kabacoff. R in action : Data analysis and graphics with R. Manning Publications Company, 2011.

[5] Irwin Miller and Maryless Miller. John E. Freund's Mathematical Statistics with Applications. Pearson Education Ltd., 2014.

[6] P.R. Halmos. Probability Theory: Independence, Interchangeability & Martingales. Springer Texts,2000.

[7] Vijay K Rohatgi and A.K Ehsanes Saleh. An Introduction to Probability and Statistics. John Wiley and Sons., 2015.

[8] Sheldon Ross. A First Course in Probability. Pearson Education Ltd., 2010.

[9] Sheldon Ross. Simulation. Academic Press, 2013.

External Links : • Wikipedia : https://en.wikipedia.org/wiki/Law_of_large_numbers

https://proofwiki.org/wiki/Khinchin's_Law

RStudio Packages : VGAM

Appendix

R Code:

For example 1:

```
# Simulation parameters
set.seed(123)
n <- 1000
X <- numeric(n)
# Generate independent, non-identical normal variables
for (i in 1:n)
{
  mu <- 1
  sigma <- sqrt(1/i)
  X[i] <- rnorm(1, mean = mu, sd = sigma)
}
# Compute running averages
running_mean <- cumsum(X) / (1:n)
# Plot the sample average
plot(1:n, running_mean, type = "l", col = "blue", lwd = 2,
     main = "WLLN: Convergence of Sample Mean",
     xlab = "n", ylab = "Sample Mean")
abline(h = 1, col = "red", lty = 2)  # True mean
legend("bottomright", legend = c("Running Mean", "True Mean = 1"),
       col = c("blue", "red"), lty = c(1, 2), bty = "n")
```

For example 2:

```
set.seed(123)
n <- 1000
X <- numeric(n)
# Generate independent, non-identical normal variables with increasing variance
for (i in 1:n)
 {
  X[i] <- rnorm(1, mean = 0, sd = i)
}
# Compute running averages
running_mean <- cumsum(X) / (1:n)
# Plot the sample average
plot(1:n, running_mean, type = "l", col = "darkgreen", lwd = 2,
     main = "Failure of WLLN: Growing Variance",
     xlab = "n", ylab = "Sample Mean")
abline(h = 0, col = "red", lty = 2)  # True mean
legend("topright", legend = c("Running Mean", "True Mean = 0"),
       col = c("darkgreen", "red"), lty = c(1, 2), bty = "n")
```

For example 3:

```
# Set number of samples
n <- 10000
# Generate normal variables with mean 0 and variance ln(i)
X <- rnorm(n, mean = 0, sd = sqrt(log(1:n)))
# Compute sample means
sample_means <- cumsum(X) / (1:n)
# Plot sample means to see convergence
plot(1:n, sample_means, type = "l", col = "blue",
     main = "Khinchin's WLLN: Sample Mean vs n",
     xlab = "n", ylab = "Sample Mean")
abline(h = 0, col = "red", lty = 2)  # True mean = 0
legend("bottomright", legend = c("Running Mean", "True Mean = 0"),
       col = c("blue", "red"), lty = c(1, 2), bty = "n")
```

For example 4:

```
# Set seed for reproducibility
set.seed(123)
# Number of terms
N <- 10000
# Generate the mean sequence
mu <- (-1)^(1:N) / (1:N)
# Standard deviation is sqrt(2) since variance = 2
sigma <- sqrt(2)
# Generate X_n ~ N(mu_n, 2)
X <- rnorm(N, mean = mu, sd = sigma)
# Compute running averages
running_avg <- cumsum(X) / (1:n)
# Plot the running average
plot(1:N, running_avg, type = "l", col = "blue",
   main = "Convergence of Sample Average (WLLN Simulation)",
   xlab = "n", ylab = "Running Average")
abline(h = 0, col = "red", lty = 2)  # True limit (0)
legend("bottomright", legend = c("Running Mean", "True Mean = 0"),
     col = c("blue", "red"), lty = c(1, 2), bty = "n")
```

For example 5:

```
# Number of samples
n <- 10000
# Generate lower and upper bounds for each Uniform distribution
a <- -1 + 1/(1:n)
b <-  1 - 1/(1:n)
# Generate Uniform samples with varying bounds
set.seed(123)  # for reproducibility
X <- runif(n, min = a, max = b)
# Compute running sample means
sample_means <- cumsum(X) / (1:n)
# Plot the running mean
plot(1:n, sample_means, type = "l", col = "blue",
     main = "Khinchin's WLLN for Uniform(-1+1/n, 1-1/n)",
     xlab = "n", ylab = "Sample Mean")
abline(h = 0, col = "red", lty = 2)
legend("topright", legend = c("Sample Mean", "Expected Value (0)"),
       col = c("blue", "red"), lty = c(1, 2), bty = "n")
```

For example 6:

```
set.seed(123)
n <- 2000
X <- numeric(n)
# Generate independent, non-identically distributed Exponential RVs
for (i in 1:n)
 {
  lambda <- 1 + 1/i
  X[i] <- rexp(1, rate = lambda)
}

# Compute running averages
running_mean <- cumsum(X) / (1:n)

# Plot running sample mean
plot(1:n, running_mean, type = "l", col = "blue", lwd = 2,
     main = "WLLN for Exponential(λ_n = 1 + 1/n)",
     xlab = "n", ylab = "Sample Mean")
abline(h = 1, col = "red", lty = 2)  # Target mean
legend("topright", legend = c("Running Mean", "Limiting Mean ≈ 1"),
       col = c("blue", "red"), lty = c(1, 2), bty = "n")
```

For example 7:

```
# Load necessary package
if (!require(VGAM)) install.packages("VGAM"); library(VGAM)
# Sample size
n <- 10000
# Function to generate Pareto(alpha) with scale = 1
rpareto1 <- function(n, alpha) {
  return((1 / runif(n))^(1 / alpha))  # Inverse transform sampling
}
# Initialize vector to store X_n values
X <- numeric(n)
# Generate independent Pareto(0.9 + 1/n) variables
set.seed(123)
for (i in 1:n) {
  alpha <- 0.9 + 1/i
  X[i] <- rpareto1(1, alpha)
}
# Compute running sample means
sample_means <- cumsum(X) / (1:n)
# Plot
plot(1:n, sample_means, type = "l", col = "red", log = "x",
     main = "Failure of WLLN for Pareto(0.9 + 1/n)",
     xlab = "n (log scale)", ylab = "Sample Mean")
abline(h = 0, col = "black", lty = 2) #True mean
legend("topleft", legend = c("Sample Mean","True mean =0"),
       col = c("red","black"), lty = c(1,2), bty = "n")
```

For example 8:

```
# Set sample size
n <- 10000

# Preallocate vector for storing gamma samples
X <- numeric(n)

# Set seed for reproducibility
set.seed(42)

# Generate independent Gamma(n, n) variables
for (i in 1:n) {
  X[i] <- rgamma(1, shape = i, rate = i)  # Gamma(i, i)
}

# Compute running sample means
sample_means <- cumsum(X) / (1:n)

# Plot sample mean convergence
plot(1:n, sample_means, type = "l", col = "blue", log = "x",
     xlab = "n (log scale)", ylab = "Sample Mean",
     main = "Khinchin's WLLN for Xn ~ Gamma(n, n)")
abline(h = 1, col = "darkgreen", lty = 2, lwd = 2)
legend("topright", legend = c("Sample Mean", "True Mean = 1"),
       col = c("blue", "darkgreen"), lty = c(1, 2), bty = "n")
```

For example 9:

```
set.seed(123)
n <- 1000
X <- numeric(n)

# Generate independent Log-Normal RVs with shrinking variance
for (i in 1:n) {
  mu <- 0
  sigma <-  sqrt(1-1/i)
  X[i] <- rlnorm(1, meanlog = mu, sdlog = sigma)
}

# Running mean
running_mean <- cumsum(X) / (1:n)

# Plot
plot(1:n, running_mean, type = "l", col = "blue", lwd = 2,
     main = "WLLN: Log-Normal with Shrinking Variance",
     xlab = "n", ylab = "Sample Mean")
abline(h = exp(0.5), col = "red", lty = 2)  # Limiting mean
legend("topright", legend = c("Running Mean", "Limiting Mean ≈ e^1/2"),
       col = c("blue", "red"), lty = c(1, 2), bty = "n")
```

For example 10:

```
# Simulation parameters
n <- 10000
sample_means <- numeric(n)
X <- numeric(n)

set.seed(123)

# Simulate scaled chi-squared(1) variables
for (i in 1:n) {
  Z_i <- rchisq(1, df = 1)
  X[i] <- Z_i / sqrt(i)
  sample_means[i] <- mean(X[1:i])
}

# Plot the sample mean
plot(1:n, sample_means, type = "l", col = "blue", lwd = 2,
     main = "Sample Mean of Scaled Chi-squared(1) Variables",
     xlab = "n", ylab = "Sample Mean")
abline(h = 0, col = "red", lty = 2)

legend("topright",
       legend = c("Sample Mean", "Limiting Value = 0"),
       col = c("blue", "red"),
       lty = c(1, 2), lwd = 2, bty = "n")
```