Formal Verification of Central Limit Theorem in HOL Theorem Prover

Thi Cam Tu Phan

A thesis submitted for the degree of

Master of Computing (Advanced)

The Australian National University

School of Computing

May 2025

Abstract

Contents

Draft ideas

0.0.1 Abstract

- What is the CLT? Why is formal verification important?
- What limitations or gaps does this work address?
- Briefly summarize the methods, results, and significance.

0.1 Introduction

Purpose: Set the stage for this thesis, establish its relevance, and provide an overview of the challenges and contributions.

0.1.1 Background of the Central Limit Theorem (CLT)

- Briefly describe the CLT's significance in probability theory and its foundational role in various fields.
- Highlight the diversity of proofs for the CLT, focusing on Lyapunov's direction as a generalization of the i.i.d. case.

0.1.2 Formal Verification of Mathematical Theorems

- Discuss the importance of formalizing theorems like the CLT for computational mathematics.
- Introduce HOL4 and its potential for advancing probability theory through rigorous formalization.

0.1.3 Challenges and Motivation

- Outline the unique challenges of formalizing Lyapunov's approach in HOL4:
- HOL4 lacks a robust library for complex analysis and probability.
- The proof depends on Taylor expansions, Big-O notation, and Lyapunov's inequality, which needed to be formalized from scratch.

- Describe the failed attempt to use moment generating functions (MGFs) and how it inspired the shift to Lyapunov's approach.

0.1.4 Thesis Contributions

- Summarize key contributions:
    1. Formalization of Lyapunov's inequality in HOL4.
    2. Development of supporting libraries for Taylor expansions, Big-O notation, and moment analysis.
    3. Complete formalization of the CLT based on Lyapunov's theorem.

0.2 Background and Related Work

Purpose: Position this work in the context of previous efforts and provide foundational concepts.

0.2.1 Overview of the Central Limit Theorem

- Explain the statement of the CLT and its generalizations (e.g., Lyapunov's and Lindeberg's approaches).

- Focus on Lyapunov's version as the basis of this proof:

$$\text{If } \frac{\Gamma_n}{s_n^3} \to 0 \text{ as } n \to \infty, \tag{1}$$

  then

$$\frac{S_n}{s_n} \xrightarrow{d} \Phi.$$

  Define key terms: Γ_n, s_n, S_n, and Lyapunov's condition.

0.2.2 Formal Proof Assistants in Probability Theory

- Discuss Isabelle/HOL and its use of characteristic functions to prove the CLT (mention Luke Serafin's work briefly here).

- Highlight HOL4's potential and this thesis's motivation to enrich its probability theory libraries.

0.2.3 Lyapunov's Direction and Proof Techniques

- Provide an informal overview of Lyapunov's proof steps:
  - Construction of normal approximations (Y_j).
  - Use of Taylor expansions to bound differences.
  - Application of Lyapunov's inequality to ensure convergence.
- Contrast Lyapunov's approach with other CLT proofs, emphasizing why it's more suitable for HOL4's existing framework.

0.3 The Formalization Process

Purpose: Detail the technical challenges and solutions in formalizing Lyapunov's proof in HOL4.

0.3.1 Overview of Lyapunov's Proof Structure

- Provide a roadmap of the proof components you formalized:
  1. Taylor expansions and approximation of test functions.
  2. Telescoping sums and bounding higher-order moments.
  3. Lyapunov's inequality and its application to ensure convergence.

0.3.2 Formalizing the Supporting Theorems in HOL4

- Taylor Expansion
- Big-O Notation
- Lyapunov's Inequality

0.3.3 Technical Challenges and Innovations

- Share how HOL4's lack of complex analysis tools required alternative approaches (e.g., using Lyapunov's proof instead of characteristic functions).
- ...

0.4 The Formal Proof of the CLT in HOL4

Purpose: Present the completed proof and validate its correctness.

0.4.1 Step-by-Step Proof

- Present the formalized proof of Lyapunov's theorem in HOL4.
- Break it into logical steps, mirroring the informal proof:
    1. Setup and normalization of S_n.
    2. Construction of Z_j and comparison with X_j.
    3. Application of Taylor expansions and bounding error terms.
    4. Application of Lyapunov's inequality to complete the proof.

0.4.2 Validation and Soundness

- Discuss how HOL4 ensures the proof's correctness and soundness.
- Highlight key intermediate results that confirm the integrity of this thesis's approach.

0.5 Results and Comparisons

Purpose: Showcase the impact of this work and compare it to related efforts.

0.5.1 Achievements in HOL4

- Summarize the completed formalization of Lyapunov's CLT proof.
- Highlight the additional libraries (e.g., Taylor expansions, Big-O notation) and their potential for future work.

0.5.2 Comparison with Isabelle/HOL

- Briefly compare this work with Luke Serafin's formalization of the CLT in Isabelle:
    - HOL4 uses Lyapunov's approach instead of characteristic functions.

- Discuss the trade-offs: HOL4's modularity and flexibility versus Isabelle's existing tools for complex analysis.

0.6 Conclusion

Purpose: Reflect on the thesis's contributions and look ahead.

0.6.1 Summary of Contributions

Restate achievements:

1. Formalized Lyapunov's theorem and the CLT in HOL4.
2. Extended HOL4's libraries for probability theory.
3. Overcame challenges related to HOL4's lack of complex analysis tools.

Also, discuss the significance of enriching HOL4 for future formalizations in probability and functional analysis.

0.6.2 Future Work

Suggest potential extensions, such as:

- Formalizing multivariate versions of the CLT.
- Extending HOL4's libraries for advanced probabilistic theorems.

0.6.3 Appendices and Supplementary Material

Code listings, etc.

0.6.4 The proof of Central Limit Theorem

0.6.4.1 Informal Proof

Informal proof of Lyapounov's theorem for a single sequence by the method of Lindeberg Chung (2000).

- **Notation and Statement**
  1. $\{X_j\}_{j=1}^n$ be a sequence of independent random variables with:
     $\mathbb{E}[X_j] = 0$,
     $\mathrm{Var}(X_j) = \sigma_j^2$ (finite variances, $\sigma_j^2 < \infty$),
     $\mathbb{E}[|X_j|^3] = \gamma_j$ (finite third moments, $\gamma_j < \infty$).
  2. Define:
     $S_n = \sum_{j=1}^n X_j$,
     $s_n^2 = \sum_{j=1}^n \sigma_j^2$ (total variance),
     $\Gamma_n = \sum_{j=1}^n \gamma_j$ (total third moment).
  3. Statement: If
     $$\frac{\Gamma_n}{s_n^3} \to 0 \quad \text{as } n \to \infty.$$
     then
     $$\frac{S_n}{s_n} \xrightarrow{d} \Phi$$
     where Φ denotes the standard normal distribution $\mathcal{N}(0, 1)$.

- **Step-by-Step Proof**
  1. **Normalize the Random Variables**
     Define the normalized variables:
     $$X_{n,j} = \frac{X_j}{s_n}, \quad \text{for } j = 1, \dots, n.$$
     The sum S_n is now expressed as:
     $$\frac{S_n}{s_n} = \sum_{j=1}^n X_{n,j}.$$
  2. **Use the Lindeberg Replacement to approximate**
     Replace each X_j with a corresponding normal random variable $Y_j \sim \mathcal{N}(0, \sigma_j^2)$, where $\{Y_j\}_{j=1}^n$ are independent and have the same mean and variance as X_j, then satisfy:
     $$\mathbb{E}[Y_j] = 0 \quad \text{and} \quad \mathrm{Var}(Y_j) = \sigma_j^2.$$
     Let all the X's and Y's be totally independent.

Since $Y_j \sim \mathcal{N}(0, \sigma_j^2)$, the normalized random variable:

$$\frac{Y_j}{s_n} \sim \mathcal{N}\left(0, \frac{\sigma_j^2}{s_n^2}\right).$$

Adding these normalized variables gives:

$$\frac{1}{s_n}\sum_{j=1}^{n} Y_j \sim \mathcal{N}\left(0, \sum_{j=1}^{n}\frac{\sigma_j^2}{s_n^2}\right).$$

Since $s_n^2 = \sum_{j=1}^{n} \sigma_j^2$, this simplifies to:

$$\frac{1}{s_n}\sum_{j=1}^{n} Y_j \sim \mathcal{N}(0, 1).$$

Now, construct the sequence Z:

$$Z_j = Y_1 + \cdots + Y_{j-1} + X_{j+1} + \cdots + X_n, \quad 1 \le j \le n.$$

Thus:

$$\begin{aligned}
Z_1 &= X_2 + X_3 + \cdots + X_n,\\
Z_2 &= Y_1 + X_3 + \cdots + X_n,\\
Z_n &= Y_1 + Y_2 + \cdots + Y_{n-1}.
\end{aligned}$$

In general, each Z_j represents a sum where:

- All variables before X_j are replaced by $Y_1, \ldots, Y_{j-1}$,
- X_j is excluded from the summation,
- All variables after X_j remain as $X_{j+1}, X_{j+2}, \ldots, X_n$.

Thus, by telescoping property, we have:

$$Y_j + Z_j = X_{j+1} + Z_{j+1}.$$

3. **Compare Distributions**

To show that $\frac{S_n}{s_n} \xrightarrow{d} \mathcal{N}(0, 1)$, we compare their expectations.

Using Test Funtions

Consider a test function f from C_B^3, the class of bounded continuous functions with three bounded derivatives.

Consider:

$$\mathbb{E}\left[f\left(\frac{S_n}{s_n}\right)\right] \quad \text{and} \quad \mathbb{E}\left[f\left(\mathcal{N}\right)\right],$$

By introducing the replacement sequence Z_j, we rewrite:

$$\mathbb{E}\left[f\left(\frac{S_n}{s_n}\right)\right] - \mathbb{E}\left[f\left(\frac{Z_n}{s_n}\right)\right].$$

By telescoping:

$$\mathbb{E}\left[f\left(\frac{S_n}{s_n}\right)\right] - \mathbb{E}\left[f\left(\frac{Z_n}{s_n}\right)\right].$$

$$= \mathbb{E}\left[f\left(\frac{X_1 + \cdots + X_n}{s_n}\right)\right] - \mathbb{E}\left[f\left(\frac{Y_1 + \cdots + Y_n}{s_n}\right)\right].$$

$$= \sum_{j=1}^{n}\left[\mathbb{E}\left[f\left(\frac{X_j + Z_j}{s_n}\right)\right] - \mathbb{E}\left[f\left(\frac{Y_j + Z_j}{s_n}\right)\right]\right].$$

4. **Expand expectations with Taylor's theorem**
   By Taylor's theorem for $f \in C_B^3$:

$$\left| f(x+y) - \left[f(x) + f'(x)y + \frac{f''(x)}{2}y^2 \right] \right| \le \frac{M}{6}|y|^3,$$

where $M = \sup_{x\in\mathbb{R}} |f^{(3)}(x)|$ is the maximum bound on the third derivative of f.

Apply this to $f(\xi + \eta)$ and take Expectations, we write:

$$\left| \mathbb{E}[f(\xi+\eta)] - \mathbb{E}[f(\xi)] - \mathbb{E}[f'(\xi)]\mathbb{E}[\eta] - \frac{1}{2}\mathbb{E}[f''(\xi)]\mathbb{E}[\eta^2] \right| \le \frac{M}{6}\mathbb{E}[|\eta|^3],$$

where ξ and η are independent random variables such that $\mathbb{E}[|\eta|^3] < \infty$.

Then, let ζ be another independent random variable, having the same mean and variance as η. Now, consider:

$$\mathbb{E}[f(\xi+\eta)] - \mathbb{E}[f(\xi+\zeta)].$$

Using Taylor's theorem again, this difference can be bounded as:

$$|\mathbb{E}[f(\xi+\eta)] - \mathbb{E}[f(\xi+\zeta)]| \le \frac{M}{6}(\mathbb{E}[|\eta|^3] + \mathbb{E}[|\zeta|^3]),$$

where the sum $\mathbb{E}[|\eta|^3] + \mathbb{E}[|\zeta|^3]$ accounts for the third absolute moments of both random variables.

Then, substitute with:

- $\xi = Z_j/s_n$,
- $\eta = X_j/s_n$,
- $\zeta = Y_j/s_n$.

We have:

$$\left|\mathbb{E}\left[f\left(\frac{Z_j}{s_n}+\frac{X_j}{s_n}\right)\right]-\mathbb{E}\left[f\left(\frac{Z_j}{s_n}+\frac{Y_j}{s_n}\right)\right]\right| \leq \frac{M}{6}\left(\mathbb{E}\left[\left|\frac{X_j}{s_n}\right|^3\right]+\mathbb{E}\left[\left|\frac{Y_j}{s_n}\right|^3\right]\right). \tag{2}$$

5. **Bound error terms using Big-O notation**

From the definition of X_j/s_n and Y_j/s_n, we know:

$$\mathbb{E}\left[\left|\frac{X_j}{s_n}\right|^3\right] = \frac{\mathbb{E}[|X_j|^3]}{s_n^3}, \quad \mathbb{E}\left[\left|\frac{Y_j}{s_n}\right|^3\right] = \frac{\mathbb{E}[|Y_j|^3]}{s_n^3}.$$

For $Y_j \sim \mathcal{N}(0, \sigma_j^2)$, the third absolute moment is:

$$\mathbb{E}[|Y_j|^3] = c\sigma_j^3,$$

where $c = \sqrt{8/\pi}$ (a constant depending on the normal distribution).
Since $\mathbb{E}[|X_j|^3] = \gamma_j$ and $\mathbb{E}[|Y_j|^3] = c\sigma_j^3$, we wrire:

$$\mathbb{E}\left[\left|\frac{X_j}{s_n}\right|^3\right]+\mathbb{E}\left[\left|\frac{Y_j}{s_n}\right|^3\right] = \frac{\gamma_j}{s_n^3}+\frac{c\sigma_j^3}{s_n^3}.$$

By summing over $j = 1, \dots, n$, (1) becomes:

$$\sum_{j=1}^{n}\left|\mathbb{E}\left[f\left(\frac{Z_j}{s_n}+\frac{X_j}{s_n}\right)\right]-\mathbb{E}\left[f\left(\frac{Z_j}{s_n}+\frac{Y_j}{s_n}\right)\right]\right| \leq \frac{M}{6}\sum_{j=1}^{n}\left(\frac{\gamma_j}{s_n^3}+\frac{c\sigma_j^3}{s_n^3}\right).$$

Lyapunov's inequality states that:

$$\sigma_j^3 \leq \gamma_j.$$

Thus:

$$\frac{\gamma_j}{s_n^3}+\frac{c\cdot\sigma_j^3}{s_n^3} \leq \frac{\gamma_j}{s_n^3}+\frac{c\cdot\gamma_j}{s_n^3} = \frac{(1+c)\cdot\gamma_j}{s_n^3}.$$

Simplify the Sum, we have:

$$\frac{M}{6}\sum_{j=1}^{n}\left(\frac{\gamma_j}{s_n^3}+\frac{c\cdot\sigma_j^3}{s_n^3}\right) \leq \frac{M}{6}\cdot\frac{1}{s_n^3}\sum_{j=1}^{n}(1+c)\gamma_j. \tag{3}$$

Let $K = \frac{M}{6}(1+c)$, and the total third absolute moment of X_j:

$$\Gamma_n = \sum_{j=1}^{n} \gamma_j (as define above).$$

We substitute these into the inequality (2):

$$\frac{M}{6} \sum_{j=1}^{n} \left(\frac{\gamma_j}{s_n^3} + \frac{c \cdot \sigma_j^3}{s_n^3} \right) \leq K \cdot \frac{\Gamma_n}{s_n^3}. \tag{4}$$

By the definition of Big-O Notation, we have

> For a function $f(n)$, we say that $f(n) = O(g(n))$ if there exist positive constants $C > 0$ and $n_0 \geq 1$ such that:
>
> $$|f(n)| \leq C|g(n)|, \quad \text{for all } n \geq n_0.$$

Here:

- $f(n) = \frac{\Gamma_n}{s_n^3}$,
- $g(n) = \frac{\Gamma_n}{s_n^3}$.

Clearly:

$$|f(n)| = \left| \frac{\Gamma_n}{s_n^3} \right| \quad \text{and} \quad |g(n)| = \left| \frac{\Gamma_n}{s_n^3} \right|.$$

Since $f(n)$ and $g(n)$ are identical, we have:

$$|f(n)| \leq 1 \cdot |g(n)| \quad \text{for all } n \geq 1.$$

Here, the constant $C = 1$. Then, the inequality $|f(n)| \leq C|g(n)|$ holds for all $n \geq 1$, so we can choose $n_0 = 1$.
Since we have found constants $C = 1$ and $n_0 = 1$ such that:

$$|f(n)| \leq C|g(n)| \quad \text{for all } n \geq n_0,$$

it follows by the definition of Big-O notation that:

$$\frac{\Gamma_n}{s_n^3} = O\left(\frac{\Gamma_n}{s_n^3} \right). \tag{5}$$

The Big O Notation property called **Multiplication by a constant** states that:
Let k be a nonzero constant. Then:

$$O(|k| \cdot g) = O(g)$$

Apply this to (4), we have:

$$K \cdot \frac{\Gamma_n}{s_n^3} = O\left(\frac{\Gamma_n}{s_n^3}\right).$$

Then (3) becomes:

$$\frac{M}{6} \sum_{j=1}^{n} \left(\frac{\gamma_j}{s_n^3} + \frac{c \cdot \sigma_j^3}{s_n^3} \right) \leq O\left(\frac{\Gamma_n}{s_n^3}\right).$$

We have thus obtained the following estimate:

$$\forall f \in C^3 : \left| \mathbb{E}\left[f\left(\frac{S_n}{s_n}\right)\right] - \mathbb{E}\left[f\left(N\right)\right] \right| \leq O\left(\frac{\Gamma_n}{s_n^3}\right).$$

6. **Convergence to Normal**

   By Lyapunov's condition:

$$\frac{\Gamma_n}{s_n^3} \to 0 \quad \text{as } n \to \infty.$$

   This directly implies that:

$$\left| \mathbb{E}\left[f\left(\frac{S_n}{s_n}\right)\right] - \mathbb{E}\left[f(N)\right] \right| \to 0.$$

   Since $\mathbb{E}[f(S_n/s_n)] \to \mathbb{E}[f(N)]$ for all bounded and continuous test functions f, by the general criterion for vague convergence (in Theorem 6.1.6 in Chung (2000)), we conclude that:

$$\frac{S_n}{s_n} \xrightarrow{d} \mathcal{N}(0, 1).$$

Chapter 1

Introduction

Formal verification has continued revolutionizing the understanding and application of foundational theorems in mathematics: the surety of accompanying rigor and its implementation in computational systems. Thus have built chances for innovative mathematical services. Probability theory is a domain with both important theoretical results and practical applications, and it has increasingly become a focus of formalization efforts. The most important from such results is the Central Limit Theorem (CLT), which carries very strong implications in theory and practice.

The CLT is perhaps the most prominent result in probability theory describing how the sum of independent random variables, if suitably normalized, approximates a standard normal distribution when the number of terms increases indefinitely. It closes the gap between random single events and predictable behavior in the aggregate, forms the basis of inferential statistics, and underpins models in a vast variety of fields such as physics, finance, and artificial intelligence. Given this importance, formal verification of the CLT using proof assistants will be an important step toward proving its correctness in automated reasoning systems.

However, formalizing CLT is far from straightforward. There exist numerous variants of this theorem, together with their numerous proofs, ranging from classical independent and identically distributed (i.i.d.) to some more general versions, such as Lyapunov's theorem **?**. Each of those proofs uses advanced mathematical tools, including characteristic functions, moment generating functions, or even higher-order moments, according to the situation. These proofs require advanced mathematical infrastructure, which is not available with all proof assistants. In particular, the case of HOL4 Slind and Norrish (2008) has found it impossible to use some of the approaches, particularly the one based on characteristic functions, because the strong complex analysis library was lacking.

Instead, this thesis follows another path: formalizing in the HOL4 proof assistant a generalized version of the CLT by Lyapunov's theorem. Lyapunov's approach is particularly appealing, as it has no direct use of either complex analysis or characteristic functions. Instead, it pays attention to moment-based convergence criteria, trying to provide its conditions for convergence using Taylor expansions, Big-O notation of error bounds, and Lyapunov's inequality. This approach aligns well with the existing mathematical infrastructure of HOL4 while simultaneously expanding its capabilities.

This project formalizes Lyapunov's proof in HOL4 and demonstrates the feasibility of verifying advanced probabilistic results in HOL4. More specifically, it has added:

- Formally defined Taylor expansions in function bounds and error terms.
- Big-O, allowing accurate representations of asymptotic behavior.
- Lyapunov's inequality, the key condition for the convergence between the CLT.

1.1 Motivation

The formalization of the Central Limit Theorem in HOL4 involved interesting challenges because of the highly mathematical nature of the theorem and the specific limitations of existing libraries in HOL4. Another difference between the proof assistants, Isabelle/HOL, which has a very mature library for complex analysis and HOL4, is that there are no built-in key concepts for working with characteristic functions and others associated with the common classical proofs of CLT.

Lack of a well-developed library has been one of the major setbacks of HOL4. Proving the central limit theorem using the more traditional kinds of proofs, such as the one formalized in Isabelle/HOL, involves the use of characteristic functions for convergence proof. Characteristic functions are essentially Fourier transforms and thereby hinge a lot on complex analysis. It would have been very difficult to follow this approach in HOL4, since the proof assistant currently lacks support for these advanced tools.

An alternative approach was explored using moment generating functions (MGFs). MGFs are much more popular for proving the CLT, wherein their specialized property is utilized to show convergence to the standard normal distribution. However, it soon became evident that formally establishing the uniqueness property of MGFs would become a serious blocker. While this property is a triviality in classical analysis, the formalization work involved was highly intricate and beyond the scope of what is currently practical within HOL4's libraries. As a result, this approach was abandoned, leading to the need for a new direction.

The breakthrough came with the adoption of Lyapunov's theorem as the basis for formalizing the CLT. Lyapunov's argument has a very direct and elementary approach to the theorem without using complex analysis or moment-generating functions but rather basing on real-valued, moment-based convergence criteria; hence, it being particularly fitted into the existing framework of HOL4. Lyapunov's proof is a milestone of abstraction in the probability theory, but it is still beautifully simple and thus met the practical and strict requirements of formal verification.

However, Lyapunov's method introduced its own set of challenges for formalization in HOL4:

- **Taylor Expansions:** The proof relies heavy on Taylor expansions to approximate test functions and bound errors. These had to be rigorously formalized in HOL4 from first principles.
- **Big-O Notation:** Reasoning about the asymptotic behavior of error terms involved the formalization of Big-O notation, which did not exist in HOL4 until now.
- **Lyapunov's Inequality:** An important part of the proof was the formalization of Lyapunov's inequality, which gives sufficient conditions for convergence. It was carefully adapted to fit into HOL4's logical framework.

Following the Lyapunov's approach, with its valuation on real techniques and elementary methods, proved most important: it served to encapsulate the CLT under the constraints of HOL4 and was entirely consistent toward building reusable mathematics libraries. Taylor series developments, Big-O notion, and Lyapunov's inequality are indispensable not only for this proof but also as a basis of further formalizations in the realm of probability theory and elsewhere.

This shift from characteristic functions and MGFs to the more concrete, moment-based approach of Lyapunov underlines flexibility and problem-solving ingenuity necessary in formal verification projects. It thus shows how advanced mathematical results can be formalized in HOL4, even when serious deficiencies in its infrastructure must first be overcome. Moreover, it emphasizes the role of Lyapunov's approach as a turning point in the history of probability theory and as one of the bases for modern formalization.

1.2 Structure of the Thesis

The rest of the thesis is organized as follows. Chapter 2 provides background on the Central Limit Theorem, its history, and the setting of Lyapunov's proof. Related work, including other formalizations, using proof assistants like Isabelle/HOL is also covered. Chapter 3 then details the formalization, discussing the challenges faced along with solutions developed, in particular the formalization of Taylor expansions, Big-O notation and Lyapunov's inequality. The key components of the proof, its completion, and its validation of the CLT are discussed in Chapter 4. Chapter 5 presents a discussion of results, and positioning of this work with respect to related efforts on formal verification. Finally, the thesis is concluded in Chapter 6 by summarizing the contributions made by this work and presenting the vision on future directions that probability theory in HOL4 may take.

This thesis not only shows that a formalization of the CLT in HOL4 is possible but also opens up further opportunities for work on the formal probability theory. It fills in some foundational gaps and provides reusable tools, which illustrates more general potential of HOL4 to make serious contributions to the fully verified mathematics corpus.

Chapter 2

Background and Related Work

The history of the Central Limit Theorem, hereinafter referred to as CLT, is quite a fascinating journey through the gradual development of probability theory. Over the years, various versions of CLT have been formulated, reflecting its adaptability and foundational importance. Each such variation addresses certain mathematical challenges and applications, therefore enriching the broader understanding of convergence in probability. In this chapter, we explore some theorem-proving tools and their methodologies to formalize the CLT, emphasizing ...

2.1 Historical Background

Let us brief the few evolvement of CLT, mainly noticing its critical features and variants developed further. The first steps towards CLT took place during the 18th century when Abraham de Moivre showed that, under summing many independent, identically distributed random variables-large number of rolled dice or coin flips-results in the normal distribution-a distribution with the shape of a bell De Moivre (1733). De Moivre laid down the very background for some approximations in the Binomial Distribution.

In 1810, Pierre-Simon Laplace expanded de Moivre's insights by proving that the sum of independent random variables converges to a normal distribution under broader conditions Laplace (1835). His work established the CLT as a universal principle for analyzing aggregate phenomena, such as population averages and measurement errors.

In the 19th century, mathematicians such as Pafnuty Chebyshev formalized the conditions of the theorem, including those of variance and expectation, so that the theorem became more precise and mathematically sound Chebyshev (1890). At the beginning of the 20th century, Lyapunov generalized CLT by introducing specific criteria-Lyapunov's condition-which explained when the theorem was applicable Lyapunov (1895). William Feller Feller (1945) later refined it to address discrete random variables.

Developments in the modern era since World War II have broadened the scope of the CLT beyond normal distributions to stable distributions and applications involving stochastic processes and high-dimensional data. These extensions illustrate how the theorem can be adapted to a wide range of probabilistic and statistical contexts. The Central Limit Theorem (CLT) has undergone significant evolution since its inception, transitioning from specific approximations to a cornerstone of modern probability theory. Among the key milestones in this journey are the foundational papers (1900/01) by Lyapunov, which marked a pivotal step toward abstraction and a more formalized understanding of the theorem. The works of Lyapunov not only reconstructed the earlier approaches but also brought in a degree of mathematical rigor consistent with the analytical standards established by Weierstrass.

Lyapunov's work differed from predecessors like Poisson and Chebyshev in that he gave an explicit, uniform bound for the difference between the distribution of the sum of random variables and the limiting normal distribution. This degree of precision met the requirement of a more precise probabilistic result, as called for by Chebyshev, and at the same time made the proof simpler. Lyapunov managed to do this by an especially elementary and at the same time subtle rewriting of Poisson's argument, including standards of rigor that had become standard in mathematics in the late 19th century.

Lyapunov's goals for the work also distinguish it. He remarks in his 1900 publication that his two main aims were to give a direct proof of the CLT - one that did not depend so heavily on the specific theories about moments developed by Chebyshev and Markov - and to relax the conditions under which the theorem was known to hold. This "direct" and "elementary" method was supposed to make the internal relationships within the theorem itself clearer and the theorem more accessible and self-contained. Indeed, Lyapunov's work represents these attributes: internally mathematically clear, abstract, yet attached to practical applicability-a sign of probability theorists of the time. His contemporaries, including von Mises, Lévy, Cramér, Khinchin, and Kolmogorov, similarly pursued abstract relationships in probability theory while eschewing purely formalistic approaches.

Yet Lyapunov himself was far from neglecting practical applicability as a criterion of mathematical achievement. The balance between working within mathematics and responding to external criteria of utility is typical of the broader development of probability theory in the modern period. This tension between purely abstract formalization and practical relevance is a defining feature of mathematical work to this day.

The generalization of CLT in Lyapunov's version is a balancing act, an extension of the classical case of independence and identical distribution. It accommodates independent random variables with only finite moments, giving way for a more flexible framework yet rigid enough to allow for a mathematical generalization. The present version forms the basis necessary for formalization in this thesis, whereby emphasis has been laid on directly obtaining real-valued methods to establish convergence under relaxed conditions.

Why Lyapunov's Version of the CLT is Chosen

Lyapunov's version of the Central Limit Theorem is particularly well-suited for formalization due to its elegant balance between generality and simplicity. Unlike earlier versions of the CLT that required independent and identically distributed (i.i.d.) random variables, Lyapunov's theorem weakens these conditions by accommodating independent random variables with potentially different variances and finite higher moments. This generalization broadens its applicability while maintaining the rigor needed for precise mathematical analysis.

One of the main reasons for selecting Lyapunov's version is that it is based on finite moments, whereas most other proofs of the CLT are based on the characteristic function approach. Though the method of characteristic functions is powerful and general, it necessarily works within the realm of complex-valued functions and makes use of Fourier transforms, which adds extra layers of abstraction and dependencies. In contrast, Lyapunov's proof stays within the realm of real-valued analysis, using explicit bounds and direct error estimates. This makes it both more accessible and more amenable to formalization in systems like HOL4, which currently has limited support for complex-valued analysis.

Moreover, Lyapunov's theorem explicitly provides a uniform bound on the error between the distribution of the normalized sum of random variables and the limiting normal distribution. The explicit control over error terms, obtained by using finite moments, is in harmony with the constructive nature of formal proofs. By employing real-valued tools such as Taylor expansions, bounding techniques, and Big-O notation, Lyapunov's proof sidesteps the intricacies of characteristic functions while retaining mathematical rigor.

In the context of this thesis, Lyapunov's is the ideal candidate for formalisation: Its direct approach relies on finite moments, focusing on real-valued methods, which perfectly suits the strengths and current capabilities of the HOL4 theorem prover. Moreover, compared to the classical i.i.d. case, Lyapunov's theorem is general and provides a stronger basis for subsequent extensions and applications.

Today, the CLT comes in a variety of variants each suited to particular situations:

- **Classical CLT:** For sums of independent, identically distributed random variables possessing finite variance and mean.
- **Generalized CLT:** When variables are weakly dependent or not identically distributed.
- **Triangular Arrays:** For sums of random variables arranged in arrays where the conditions vary across rows.
- **Local and Integral Versions:** While the local CLT concerns pointwise probabilities,

the integral version deals with cumulative distributions.

From about 1810 to 1935, most of the efforts were devoted to proving the CLT for sums of independent random variables, with more recent generalizations involving only weakly dependent variables. Modern formulations neatly distinguish between normed sums, triangular arrays, and local versus integral theorems.

Chapter 3

Preliminaries

Preliminaries 123

Chapter 4

Central Limit Theorem

asldfjkldasjflkdfjldsa;f

Chapter 5

Future Work

Future Work

Chapter 6

Conclusion

asdfsdfasdf

asdasasfdfdf

References

Yves Bertot and Pierre Castéran. *Interactive theorem proving and program development: Coq'Art: the calculus of inductive constructions*. Springer Science & Business Media, 2013.

Manfred Broy, Bengt Jonsson, J-P Katoen, Martin Leucker, and Alexander Pretschner. Model-based testing of reactive systems. In *Volume 3472 of Springer LNCS*. Springer, 2005.

Jerry R Burch, Edmund M Clarke, Kenneth L McMillan, David L Dill, and Lain-Jinn Hwang. Symbolic model checking: 1020 states and beyond. *Information and computation*, 98 (2):142–170, 1992.

Pafnutii Lvovich Chebyshev. Sur deux théorèmes relatifs aux probabilités. *Acta math*, 14 (1):305–315, 1890.

Kai Lai Chung. *A course in probability theory*. Elsevier, 2000.

Edmund M Clarke, Thomas A Henzinger, Helmut Veith, Roderick Bloem, et al. *Handbook of model checking*, volume 10. Springer, 2018.

Abraham De Moivre. *Approximatio ad summam terminorum binomii (a+ b) n in seriem expansi*. 1733.

Willy Feller. The fundamental limit theorems in probability. *Bulletin of the American Mathematical Society*, 51:800–832, 1945. Reprinted in Adams 2009, pp. 80–113.

Hans Fischer. *A history of the central limit theorem: from classical to modern probability theory*, volume 4. Springer, 2011.

John Harrison. Hol light, 2015. URL `http://www.cl.cam.ac.uk/~jrh13/hol-light/`. Accessed: 2024-12-01.

Pierre Simon Laplace. *Oeuvres complètes de Laplace*. Gautier-Villars, 1835.

Aleksandr Mikhailovich Lyapunov. Pafnutii lvovich chebyshev, 1895.

Sheldon Ross. *First Course in Probability, A*. Pearson Higher Ed, 2019.

Konrad Slind and Michael Norrish. A brief overview of hol4. In *International Conference on Theorem Proving in Higher Order Logics*, pages 28–32. Springer, 2008.

Isabelle Team. Isabelle/hol, 2015. URL `https://isabelle.in.tum.de/`. Accessed: 2024-12-01.

Index