Lehigh University
Lehigh Preserve

Theses and Dissertations

2013

The Central Limit Theorem and the Estimation of the Concentration of Measure for Fractional Brownian Motion

Patricia Mehron Garmirian
Lehigh University

Follow this and additional works at: http://preserve.lehigh.edu/etd

Part of the Mathematics Commons

Recommended Citation
Garmirian, Patricia Mehron, "The Central Limit Theorem and the Estimation of the Concentration of Measure for Fractional Brownian Motion" (2013). *Theses and Dissertations.* Paper 1491.

This Dissertation is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Lehigh Preserve. For more information, please contact preserve@lehigh.edu.

The Central Limit Theorem and the Estimation of the Concentration of Measure for Fractional Brownian Motion

by

Patricia Mehron Garmirian

A Dissertation
Presented to the Graduate Committee
of Lehigh University
in Candidacy for the Degree of
Doctor of Philosophy
in
Mathematics

Lehigh University
September 2013

Copyright
Patricia Garmirian

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Patricia Mehron Garmirian

The Central Limit Theorem and the Estimation of the Concentration of Measure for Fractional Brownian Motion

Date

Vladimir Dobric, Dissertation Director, Chair

Accepted Date

Committee Members

Evan Fisher

Terrence Napier

Lee Stanley

Acknowledgments

I thank my parents, James and Barbara Garmirian, and my brother, David, for their constant love and support throughout the years. Without their consistent encouragement, this dissertation would not be possible. My family is everything to me.

I thank my advisor Vladimir Dobric for his endless patience, kindness, wisdom, and intelligence that made this dissertation possible. Having worked with him, I have not only grown as a mathematician but also as a human being. I also thank Lee Stanley for inviting me to join his research meetings with Professor Dobric and for his guidance and teachings over the years. I thank my committee members, Lee Stanley, Terrence Napier, and Evan Fisher for their attentive involvement in my dissertation and for their helpful questions and suggestions.

Finally, I thank the math professors at Denison University for instilling me with a love of math that I have maintained throughout my graduate studies. I thank Professors Bonar, Gutman, Hutson, Ludwig, and Neal for teaching me mathematics and preparing me for my studies at Lehigh. I especially thank Professor Matthew Neal for introducing me to mathematical research through our work in functional analysis.

Contents

Abstract

The principal result of Chapter 1 is a new, direct and elementary proof of the general Central Limit Theorem (CLT). Two important stepping-stones are, first, a new, similarly direct and elementary proof of the CLT for Bernoulli random variables defined on [0,1]; this was initially proved by Bernoulli in the 1700's. The second important stepping-stone is a new result for Bernstein polynomials of continuous functions. Bernstein polynomials are a fundamental object of mathematical analysis. It is well known that Bernstein polynomials of a continuous function on intervals $[0, b_n]$ when n tends to infinity return the value of the function for an appropriate rate of b_n, but uniform convergence is sacrificed. Nothing was known for the symmetric interval $[-b_n, b_n]$. We have proven that for these intervals the limit does not recover the function but rather its integral with respect to Gaussian measure. The extension to our direct proof of the of the general CLT involves a new and surprising connection between the CLT and the Haar basis on [0, 1]: the i.i.d. sequence of random variable is transformed to a sequence defined on [0,1] and the random variables in the transformed sequence are then expanded with respect to the Haar basis.

Our work on the estimation of the concentration of measure for fractional Brownian motion requires finding the intersections of ellipsoidal and spherical shells for Gaussian measure in $\mathbb{R}^N$. Gaussian measure is concentrated on a small shell of a sphere of radius the square root of N. We want to determine how large this shell must be to include the majority of the Gaussian measure. This result determines the rate of convergence of averages of squares for fractional Brownian increments. It requires understanding the spectrum of the covariance operator as a function of dimension N and the Hurst index. To help understand the spectrum, we compute the exact rate of the largest eigenvalue of this operator.

Introduction

In this dissertation, we will present our new proof of the Central Limit Theorem (CLT) and our computation of the confidence intervals for fractional Brownian motion. In chapter 1, we will discuss our proof of the CLT. While the standard proof of today makes use of the Levy Continuity Theorem, our proof avoids this theorem to provide a direct proof of the CLT. We accomplish this by expanding our random variables using the Haar wavelet basis. In section 1, we will discuss the various definitions for weak convergence and why weak-$*$ convergence in the dual space of bounded, regular, finitely additive measures is the most natural definition. In section 2, we will discuss the history of the theorem and how it relates to our new proof. In section 3, we will discuss Bernstein polynomials, a fundamental object of mathematical analysis. In section 4, we will give a new elementary proof of the CLT for Bernoulli random variables. This proof not only establishes a new CLT result but also provides a new result for Bernstein polynomials. In section 5, we will dicuss the Haar wavelet basis and how we use this basis to give a new proof of the CLT.

In chapter 2, we will present our computation of the confidence intervals for fractional Brownian motion (fBm). In section 1, we will provide definitions and explain the applications of Brownian motion and fBm. In section 2, we will present our computation for the confidence intervals for fBm which uses ergodic theory and Jensen's inequality. This computation requires knowledge of the spectrum for the covariance operator for fBm increments. In section 3, we will compute the largest eigenvalue for the covariance matrix. This gives us an estimation for the concentration of measure for fBm.

Chapter 1

The Central Limit Theorem

1.1 Background

We have succeeded in giving a new proof of the Central Limit Theorem. There are four main types of convergence for sequences of random variables: almost sure convergence, convergence in probability, L^p convergence, and weak convergence. For definitions of these types of convergence, consult Appendix A. Weak convergence is the type of convergence required for the Central Limit Theorem. A sequence of real-valued random variables (X_n) on a probability space $(\Omega, \mathcal{F}, P)$ is said to converge weakly, denoted by "$\Rightarrow$", to a random variable X provided that for each bounded, continuous function $f : \mathbb{R} \to \mathbb{R}$, we have

$$\lim_{n\to\infty} E_P(f(X_n)) = E_P(f(X)).$$

By the Helly-Bray Theorem, this statement is equivalent to the following definition: For each $x \in \mathbb{R}$,

$$\lim_{n\to\infty} P(X_n \leq x) = P(X \leq x).$$

By Levy's Continuity Theorem, weak convergence is equivalent to the following:

$$\lim_{n\to\infty} E_P(e^{itX_n}) = E_P(e^{itX})$$

for each $t \in \mathbb{R}$. The functions $\phi(t) = E_P(e^{itX})$, $\phi_n(t) = E_P(e^{itX_n})$ are known as the characteristic functions for X, X_n, respectively. We will see that the first formulation type involving bounded, continuous functions is the most natural definition for weak convergence. Weak convergence of random variables is discussed in more detail in [2] and [3].

Recall that a Banach space X is a complete normed vector space. The dual space X^* of X is the space of all continuous linear functionals on X. That is, X^* is the space of all continuous linear functions $\phi : X \to \mathbb{R}$. X^* is also a Banach space with the norm given by

$$\|\phi\| = \sup_{\|x\| \leq 1} |\phi(x)|.$$

We may subsequently consider the dual space X^{**} of X^*, known as the second dual of X. We will construct an imbedding $\kappa : X \hookrightarrow X^{**}$. For each $x \in X$, let $\kappa(x) \in X^{**}$ be the linear functional on X^* given by

$$\kappa(x)(\phi) = \phi(x)$$

for each $\phi \in X^*$. Then, $\kappa : X \hookrightarrow X^{**}$ is an injective linear map. Hence, we can think of X as a subset of X^{**} as $\kappa(X) \subseteq X^{**}$.

A sequence of elements $\{x_n\} \subseteq X$ is said to converge in the weak topology to $x \in X$ provided that

$$\lim_{n \to \infty} \phi(x_n) = \phi(x)$$

for each $\phi \in X^*$. A sequence of elements $\{\phi_n\} \subseteq X^*$ is said to converge in the weak$-^*$ topology to $\phi \in X^*$ provided that

$$\lim_{n \to \infty} \phi_n(x) = \phi(x)$$

for each $x \in X$. The weak$-^*$ topology is the coarsest topology on X^* in which the maps

$$\kappa(x) : X^* \to \mathbb{R}$$

are continuous.

Let S be a separable, complete metric space (Polish space), and consider $C_b(S)$, the space of all bounded, continuous functions on S. Then, by the Riesz Representation Theorem, any linear functional L on $C_b(S)$ has the form

$$L(f) = \int_S f d\mu$$

for some regular, bounded, finitely additive measure μ defined on the Borel σ-algebra $\mathcal{B}_S$. Equipped with the norm given by total variation of measure, the space of all regular, bounded, finitely additive measures on S, rba(S), is a Banach space with $(C_b(S))^*$ =rba(S). Thus, for a sequence of measures $\{\mu_n\}$ from rba(S) to converge to a measure $\mu \in$ rba(S) in the weak$-^*$ topology, it is required that

$$\lim_{n\to\infty} \mu_n(f) = \mu(f)$$

for each $f \in C_b(S)$. Thus, by the Riesz Representation Theorem, we have

$$\lim_{n\to\infty} \int_S f d_{\mu_n} = \int_S f d\mu$$

for each $f \in C_b(S)$. Letting $S = \mathbb{R}$, $\mu_n = PX_n^{-1}$, and $\mu = PX^{-1}$, we arrive at the definition for weak convergence. Functional analysis is discussed in more detail in [6].

The Central Limit Theorem (CLT), the second pearl of probability theory, states that if (X_i) is a sequence of independent identically distributed random variables with $E(X_1) = \mu$ and $\mathrm{var}(X_1) = \sigma^2$, then

$$\frac{X_1 + ... + X_n - n\mu}{\sqrt{n}\sigma} \Rightarrow N(0,1).$$

Here, $N(0,1)$ denotes the normal distribution with mean 0 and variance 1. This distribution is also known as the famous "bell curve" from statistics. A random variable $Y \overset{d}{=} N(0,1)$ provided that

$$P(a \leq Y \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

for each $a, b \in \mathbb{R}$ with $-\infty < a \leq b < \infty$ where $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ is the density function for $N(0,1)$. The current standard proof of the CLT establishes the convergence of the characteristic functions for our sequence of random variables. The CLT is discussed in more detail in [2] and [3].

1.2 History

The CLT is the result of contributions made by such famous mathematicians, interested in the distribution of sums of independent random variables, as Laplace, Poisson, Dirichlet, Cauchy, Chebyshev, Markov, Feller, and Levy, among others. In this section, I will discuss the contributions made by these mathematicians to the CLT from the 1770s until the 1930s. The following historical sketch is based on [4].

Laplace was the first mathematician to break significant ground on the CLT. According to H. Fischer, before Laplace and his successors, applications of probability theory mainly involved moral problems. Formulas existed for computing probabilities based on a large number of trials, but they were too complicated for numerical calculations. In 1774, Laplace made his first efforts towards proving the CLT by developing useful methods for approximating the probabilities of sums of independent random variables. In 1810, he made significant progess through the use of generating functions and the clever substitution $t = e^{ix}$. Laplace's special case of the CLT was the result of forty years of effort. For more details on Laplace's work on the CLT, consult Appendix B.

Laplace never proved the general CLT that we use today. Instead, he considered the approximate probabilities involving linear combinations of observed errors. His most general version of the CLT is the following: Let $\epsilon_1, ..., \epsilon_n$ be independent observation errors with mean μ and variance σ^2. Let $\lambda_1, ..., \lambda_n$ be constant multipliers and $a > 0$. Then, we have

$$P\left(|\sum_{j=1}^{n}\lambda_j(\epsilon_j - \mu)| \leq a\sqrt{\sum_{j=1}^{n}\lambda_j^2}\right) \approx \frac{2}{\sigma\sqrt{2\pi}}\int_0^a e^{-\frac{x^2}{2\sigma^2}}\,dx.$$

Although Laplace never proved today's general CLT, he did introduce several new ideas that inspired the work of his successors, including Poisson. However, while Laplace and Poisson agreed on the study of probability in a classical sense, they differed in its applications to moral problems. While Laplace exercised caution with regards to these applications, Poisson believed that the laws of mathematics had a direct connection to the physical world. He strived to use precise mathematical analysis to solve real world problems. Through his work, Poisson established a formula for the probability that a sum of random variables is within given limits. He then produced a counterexample to this formula, which led to the reworking of his assumptions. For more details on Possion's work on the CLT, consult Appendix C.

In the nineteenth century, many changes in mathematics occurred. The field of probability was criticized for its use in human decision making, such as court trials. Dirichlet's main interest in mathematics was in the discussion of analytical problems rather than in applications. Like Laplace, Dirichlet also worked with the gamma function.

$$\Gamma(s+1) = M\int_{-s}^{\infty} e^{-z}\left(1+\frac{z}{s}\right)^{s} dz = M\int_{-s}^{\infty} e^{-t^2}\frac{dz}{dt}dt.$$

Setting

$$e^{-z}\left(1+\frac{z}{s}\right)^{s} = e^{-t^2},$$

he differentiated both sides to obtain

$$z\frac{dz}{dt} = 2t(s+z).$$

Using the expansion

$$z = k_1 t + k_2 t^2 + ...,$$

Dirichlet established the recursive formulas

$$k_1 = \sqrt{2s}, k_n = \frac{2k_{n-1}}{(n+1)k_1} - \frac{1}{2k_1}\sum_{i=2}^{n-1} k_i k_{n+1-i}$$

to obtain

$$\Gamma(s+1) = s^{s+\frac{1}{2}}e^{-s}\sqrt{2\pi}\left(1+\sum_{n\geq 1}\frac{1\cdot 3\cdot 5...(2n+1)a_{2n+1}}{s^n}\right),$$

where

$$a_i = 2^{1-i}(\sqrt{2s})^{i-2}k_i.$$

He set

$$\int_{-n}^{\infty} e^{-z}(1+\frac{z}{n})^n dz = \int_{-n}^{\infty} y dz = \Gamma(n+1)e^n n^{-n},$$

where
$$\int_{-n}^{-n^m} ydz + \int_{-n^m}^{n^m} ydz + \int_{n^m}^{\infty} ydz = I_1 + I_2 + I_3,$$
and $\frac{1}{3} < m < \frac{2}{3}$. Then, he showed that $I_1, I_3 \to 0$, and
$$\frac{I_2}{\sqrt{2n}} \to \int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}.$$
Thus, Dirichlet obtained a result for $\Gamma(n+1)$ used to approximate the normal distribution.

Cauchy made an important contribution to the CLT by introducing the concept of the characteristic function which is used today. This development resulted from Cauchy's analysis of the interpolation of random errors. In the final paper of eight which he published on the CLT, Cauchy discussed the approximate normal distribution of linear combinations of random errors. This argument is similar to the one made by Dirichlet, and his method is still used today.

In studying the interpolation of observational errors, Cauchy used the "fonction auxiliaire" now known as the characteristic function for a random variable. With the error in $[\kappa_1, \kappa_2]$ given by the function $g(x)$, we have "fonction auxiliaire" given by
$$\phi(x) = \int_{\kappa_1}^{\kappa_2} e^{-izx} g(z) dz.$$
Cauchy's version of the CLT provides upper bounds for the error of the normal approximation to the distribution of
$$\sum_{j=1}^{n} \lambda_j \epsilon_j$$
where (ϵ_j) are idependent identically distributed errors. He assumed that (λ_j) have order $O(\frac{1}{n})$ and $\sum \lambda_j^2 := \Lambda$ has the order of $O(\frac{1}{n})$. Then, he established
$$|P\left(-v \le \sum_{j=1}^{n} \lambda_j \epsilon_j \le v\right) - \int_0^{\frac{v}{2\sqrt{c\Lambda}}} e^{-\theta^2} d\theta| \le C_1(n) + C_2(n,v) + C_3(n)$$

for sufficiently large n, with explicit formulas for C_1, C_2, C_3.

Chebyshev and his student Markov made contributions to the CLT through their method of moments. An enthusiastic teacher, Chebyshev founded the St. Petersburg school which Markov attended. The method of moments involves finding properties of monotonically increasing functions $\mu \geq 0$ defined on $[a, b]$ by knowing its moments

$$M_0 := \int_{x\in[a,b]} d\mu(x),\ \ M_1 := \int_{x\in[a,b]} x d\mu(x), ...,\ \ M_n := \int_{x\in[a,b]} x^n d\mu(x).$$

In 1887, Chebyshev introduced the following version of the CLT (in the terminology of today): Let u_i be a sequence of independent random variables (quantities) with zero expectation, nonnegative densities, ϕ_i, with finite moments of arbitrarily high order. Assume that, for each order, an upper and lower bound of the moments exist, uniformly in i but not in n. Then, for any $t < t' \in \mathbb{R}$, we have

$$\lim_{n\to n} P\left(t \leq \frac{\sum_{i=1}^{n} u_i}{\sqrt{2\sum_{i=0}^{n} Eu_i^2}} \leq t'\right) = \frac{1}{\sqrt{\pi}} \int_t^{t'} e^{-x^2} dx.$$

In 1898, Markov published his version of the CLT with the following assumptions: Let $u_1,\ u_2, ...$ be "independent quantities" which satisfy the following conditions: $Eu_k = 0$ for each k. For all natural numbers $m \geq 2$ there exists a constant C_m such that $|Eu_k^m| < C_m$ for all $k \in \mathbb{N}$. Eu_k^2 "does not get infinitely small, if k grows indefinitely." Then,

$$\lim_{n\to\infty} P\left(\alpha\sqrt{2\sum_{i=0}^{n} Eu_i^2} \leq \sum_{i=1}^{n} u_i \leq \beta\sqrt{2\sum_{i=0}^{n} Eu_i^2}\right) = \frac{1}{\sqrt{\pi}} \int_\alpha^\beta e^{-x^2} dx$$

for $\alpha < \beta \in R$. Notice that the conclusions for Chebyshev's and Markov's versions of the CLT are essentially identical.

In the twentieth century, the field of probability evolved to become more rigorous. At the second International Congress of Mathematics in Paris in 1900, Hilbert proposed the axiomatization of the applied sciences as one of his 23 problems. This problem required "to treat in the same manner, by means of axioms, those physical sciences in which mathematics plays an important part; in the first rank are the theory of probability and mechanics." This proposition set probability on a path towards becoming axiomatic, due to Kolmogorov in 1933, and following precise mathematical analysis.

This goal of axiomatizing probability was achieved by Kolmogorov in 1933. [4] A σ-algebra on the set Ω is a collection of subsets of Ω which contains the empty set and is closed under complements and countable unions. A probability measure on a set Ω is a set function on a σ-algebra $\mathcal{F}$ which satisfies the following axioms:

$a)$ $0 \leq P(A) \leq 1$ for all $A \in \mathcal{F}$

$b)$ $P(\emptyset) = 0, \; P(\Omega) = 1$

$c)$ For any sequence $A_1, A_2, ...$ of disjoint sets in $\mathcal{F}$,

$$P\left(\cup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k).$$

These axioms provided a universal setting for later work on the CLT. Levy then proved his version of the CLT after first proving some preliminary results about characteristic functions which, in turn, he used in proving his version of the CLT. Levy was introduced to characteristic functions while reading the work of Poincare. Poincare defined "fonctions characteristiques" to be functions of the form

$$f(\alpha) = \sum p(x) e^{\alpha x}$$

for discrete quantities whose values x occur with probability $p(x)$, and

$$f(\alpha) = \int_{-\infty}^{\infty} \phi(x) e^{\alpha x} dx$$

for continuous quantities with density ϕ. Levy then proved the following theorems about characteristic functions: Let $\phi(z)$ denote the characteristic function for distribution function $F(x)$ given by

$$\phi(z) = \int_{-\infty}^{\infty} e^{izx} dF(x).$$

Theorem 1: If for $\lambda \to \lambda_0$ the laws L_λ tend to the limit law $\mathcal{L}$ with characteristic function ω, then $\phi_\lambda(z)$ also tends to $\omega(z)$ uniformly in each compact interval of z-values.
Theorem 2: If ω is a characteristic function such that $\lim_{\lambda \to \lambda_0} \phi_\lambda(z) = \omega(z)$ uniformly in each compact interval of z-values, then L_λ tends to the probability law $\mathcal{L}$ which belongs to ω.

Levy and Feller each proved a version of the CLT. Feller's version of the CLT can be stated as follows: Let (X_k) be a sequence of independent random variables with distributions V_k all having median 0. Then, there exist sequences $(a_n > 0)$ and (b_k) of real numbers such that

$$P\left(\frac{1}{a_n}\sum_{k=1}^{n}(X_n - b_k) \leq x\right) \to \Phi(x).$$

Further,

$$\max_{1 \leq k \leq n} P(|X_k - b_k| > \epsilon a_n) \to 0 \text{ for each } \epsilon > 0$$

as $n \to \infty$ if and only if

$$\text{for each } \delta > 0, \text{ for each } \eta > 0, \text{ there exists } n(\delta, \eta) \text{ such that for each } n \geq n(\delta, \eta):$$

$$\frac{p_n^2(\delta)}{\sum_{k=1}^{n}\int_{|x| \leq p_n(\delta)} x^2 dV_k(x)} < \eta$$

where $p_n(\delta) = \min\{r \in \mathbb{R}_0^+ : P(|X_k| > r) \leq \delta\}$.

Levy's version of the CLT can be stated as follows: Let L_n be the dispersion of $\sum_{k=1}^{n} X_k$ having a fixed probability $\gamma \in (0,1)$. Then, there exist sequences $(a_n > 0)$ and (b_k) of real numbers such that

$$P\left(\frac{1}{a_n}\sum_{k=1}^{n}(X_n - b_k) \leq x\right) \to \Phi(x).$$

Further,

$$\max_{1 \leq k \leq n} P(|X_k| > \epsilon L_n) \to 0 \text{ for each } \epsilon > 0$$

as $n \to \infty$ if and only if for each $\delta > 0$, for each $\eta > 0$, there exists $n(\delta, \eta)$ such that for each $n \geq n(\delta, \eta)$ there exists $X(n) > 0$:

$$\frac{X^2(n)}{\sum_{k=1}^{n}\left(\int_{|x|\le X(n)} x^2 dV_k(x) - \left(\int_{|x|\le X(n)} x dV_k(x)\right)^2\right)} < \eta$$

and

$$\sum_{k=1}^{n} P(|X_k| > X(n)) < \delta.$$

Efforts to give new proofs of the CLT have been made since the proofs of Levy and Feller. In [1], Barron shows that the density function $f_n(x)$ a normalized sum of i.i.d. random variables converges to the normal density $\phi(x)$ in the sense of relative entropy: $\int f_n \ln(f_n)/\phi \to 0$ provided that relative entropy is finite for some n. In [7], Bahr analyzes the convergence of moments of normalized sums of i.i.d. random variables towards corresponding moments of the normal distribution.

1.3 Bernstein polynomials

The Bernstein polynomial of order n of the function $f(x)$ defined on the closed interval $[0, 1]$ is given by

$$B_n(x) = B_n^f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}.$$

For f continuous on $[0, 1]$,

$$\lim_{n\to\infty} B_n(x) = f(x)$$

uniformly in x. Bernstein first introduced this set of polynomials to provide a simple proof of the Weierstrauss Approximation Theorem. As we may transform the interval $[a, b]$ into $[0, 1]$, the result holds for a function f on any closed, bounded interval. The so-called ”singular operators” provide other means for approximating a generating function $f(x)$. The best known singular operator is the Dirichlet integral

$$s_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\frac{\sin(n+\frac{1}{2})(t-x)}{2\sin\frac{1}{2}(t-x)} dt$$

which represents the partial sums $s_n(x)$ of the Fourier series of the function $f(x)$ integrable on $[-\pi, \pi]$.

We can also define Bernstein polynomials for functions on unbounded intervals. The Bernstein polynomial of order n defined on the interval $(0, b)$ is found by making the substitution $y = \frac{x}{b}$ in the polynomial $B_n^\phi(y)$ of the function $\phi(y) = f(by)$, $0 \leq y \leq 1$. Thus, we obtain the polynomial

$$B_n(x) = B_n^f(x; b) = \sum_{k=0}^{n} f\left(\frac{bk}{n}\right)\binom{n}{k}\left(\frac{x}{b}\right)^k\left(1 - \frac{x}{b}\right)^{n-k}.$$

By letting $b = b_n$, a function of n, we may consider a function f on the unbounded interval $(0, \infty)$. As with functions on $[a, b]$, we would like for $B_n(x; b_n) \to f(x)$ to hold with minimal assumptions on f. It is true that this relation is preserved for $b_n = o(n)$. As the example of the function $f(x) = x^2$ with $B_n(x; b_n) = \left(1 - \frac{1}{n}\right)x^2 + b_n\frac{x}{n}$ provides a counterexample, this condition is also necessary. The material on Bernstein polynomials is develped in [5].

The next question to naturally arise is whether we can extend this relation to functions on unbounded, symmetric intervals. We start by considering the Bernstein polynomial of order n defined on the interval $(-b, b)$. This polynomial is found by making the substitution $y = \frac{x+b}{2b}$ in the polynomial $B_n^\phi(y)$ of the function $\phi(y) = f(b(2y - 1))$, $0 \leq y \leq 1$. Thus, we obtain the polynomial

$$B_n(x) = B_n^f(x; -b, b) = \sum_{k=0}^{n} f\left(2b\left(\frac{k}{n}\right) - b\right)\binom{n}{k}\left(\frac{x+b}{2b}\right)^k\left(1 - \left(\frac{x+b}{2b}\right)\right)^{n-k}.$$

For $b = \sqrt{n}$, we have

$$\begin{aligned} f\left(2b\left(\frac{k}{n}\right) - b\right) &= f\left(\frac{2\sqrt{n}k}{n} - \sqrt{n}\right) \\ &= f\left(\frac{2k}{\sqrt{n}} - \frac{n}{\sqrt{n}}\right) \\ &= f\left(\frac{2k-n}{\sqrt{n}}\right). \end{aligned}$$

In the next section, we show that for any bounded, continuous function $f : \mathbb{R} \to \mathbb{R}$, we have

$$B_n^f(0; -\sqrt{n}, \sqrt{n}) = \sum_{k=0}^{n} f\left(\frac{2k-n}{\sqrt{n}}\right)\binom{n}{k}\left(\frac{1}{2}\right)^n \to \int_{-\infty}^{\infty} f(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx.$$

It is a simple exercise to extend from $x = 0$ to any $0 \le x \le 1$.

1.4 A new proof of the CLT for Bernoulli random variables

In this section, we will present our new elementary proof of the CLT for Bernoulli random variables defined on $[0, 1]$. Bernoulli is the first mathematician to consider the CLT for Bernoulli random variables.

Let (X_i) be an i.i.d. sequence of random variables with $P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}$. Then, $EX_1 = 0$ and $\text{var}X_1 = 1$. We will verify the Central Limit Theorem by showing that

$$\frac{X_1 + ... + X_n}{\sqrt{n}} \Rightarrow N(0, 1).$$

We will prove this by showing that for any bounded continuous function $f : \mathbb{R} \to \mathbb{R}$, we have

$$\lim_{n\to\infty} E_P\left(f\left(\frac{X_1 + ... + X_n}{\sqrt{n}}\right)\right) = E_P(f(Y))$$

where $Y \overset{d}{=} N(0, 1)$. To show this, we will first compute

$$E_P\left(f\left(\frac{X_1 + ... + X_n}{\sqrt{n}}\right)\right).$$

We have n i.i.d. random variables each having the values 1 and (-1) with probability $\frac{1}{2}$. If we consider the sum of these random variables, then for any point in the sample space, we have k of the random variables equal to 1 and the other $(n - k)$ of them equal to (-1) for some $0 \le k \le n$. Summing up the random variables, we have the values $k - (n - k) = 2k - n$ for $0 \le k \le n$. By independence, we multiply the probabilities for each value of X_i, $1 \leq i \leq n$, to obtain the probability of $\left(\frac{1}{2}\right)^n$ for every combination of $X_i = \pm 1$, $1 \leq i \leq n$. As the sum of the random variables is equal for every combination of k 1's and $(n-k)$ (-1)'s , the value $\frac{2k-n}{\sqrt{n}}$ has probability $\left(\frac{1}{2}\right)^n \binom{n}{k}$. Therefore, we have

$$E_P\left(f\left(\frac{X_1 + ... + X_n}{\sqrt{n}}\right)\right) = \sum_{k=0}^{n} \left(\frac{1}{2}\right)^n \binom{n}{k} f\left(\frac{2k-n}{\sqrt{n}}\right).$$

In the previous section, we discussed Bernstein polynomials on symmetric intervals. The right hand side of the previous displayed equation is the Bernstein polynomial $B_n^f(0; -\sqrt{n}, \sqrt{n})$. Since $Y \stackrel{d}{=} N(0,1)$, then we have $E_P(f(Y)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{\frac{-x^2}{2}} dx$. Therefore, we will show that

$$\lim_{n \to \infty} \sum_{k=0}^{n} \left(\frac{1}{2}\right)^n \binom{n}{k} f\left(\frac{2k-n}{\sqrt{n}}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{\frac{-x^2}{2}} dx.$$

Let $\epsilon > 0$. Choose b such that $\frac{1}{b} < \frac{\epsilon}{6}$ and $\frac{1}{\sqrt{2\pi}} \int_b^{\infty} e^{\frac{-x^2}{2}} dx < \frac{\epsilon}{12\|f\|_\infty}$. By Chebyshev's Inequality, we have

$$\sum_{|\frac{n}{2}-k| > \frac{b\sqrt{n}}{2}} (\frac{1}{2})^n \binom{n}{k} \leq \frac{\sqrt{\frac{1}{4}n}}{\frac{b\sqrt{n}}{2}} = \frac{1}{b} < \frac{\epsilon}{6}.$$

Since $g(x) = f(x)\frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$ is continuous on $\mathbb{R}$, then $g(x)$ is uniformly continuous on $[-b, b]$. Therefore,

$$\lim_{n \to \infty} \sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}} g\left(\frac{2k-n}{\sqrt{n}}\right) \chi_{[\frac{2k-n}{\sqrt{n}}, \frac{2(k+1)-n}{\sqrt{n}})}(x) = g(x)\chi_{[-b,b)}(x).$$

By the Dominated Convergence Theorem,

$$\lim_{n \to \infty} \sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}} g\left(\frac{2k-n}{\sqrt{n}}\right) \frac{2}{\sqrt{n}} = \int_{-b}^{b} g(x) dx.$$

Therefore, there exists an $N \in \mathbb{N}$ such that

$$|\sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}} g\left(\frac{2k-n}{\sqrt{n}}\right) \frac{2}{\sqrt{n}} - \int_{-b}^{b} g(x) dx| < \frac{\epsilon}{6}$$

for all $n \geq N$. Let

$$A_n := |\sum_{k=0}^{n} (\frac{1}{2})^n \binom{n}{k} f\left(\frac{2k-n}{\sqrt{n}}\right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{\frac{-x^2}{2}} dx|.$$

Therefore, by cutting off the tails of our integral, for all $n \geq N$,

$$A_n \leq |\sum_{k=0}^{n} (\frac{1}{2})^n \binom{n}{k} f\left(\frac{2k-n}{\sqrt{n}}\right) - \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} f(x) e^{\frac{-x^2}{2}} dx| + \frac{\epsilon}{6}.$$

Using Chebyshev's inequality,

$$A_n \leq |\sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}} (\frac{1}{2})^n \binom{n}{k} f\left(\frac{2k-n}{\sqrt{n}}\right) - \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} f(x) e^{\frac{-x^2}{2}} dx| + \frac{\epsilon}{3}.$$

For $n \geq N$,

$$A_n \leq |\sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}} (\frac{1}{2})^n \binom{n}{k} f\left(\frac{2k-n}{\sqrt{n}}\right) - \sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}} f(\frac{2k-n}{\sqrt{n}}) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{2k-n}{\sqrt{n}})^2} \frac{2}{\sqrt{n}}| + \frac{\epsilon}{2}.$$

Since f is a bounded function,

$$A_n \leq \|f\|_\infty \sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}} |(\frac{1}{2})^n \binom{n}{k} - \frac{1}{\sqrt{2\pi}} e^{-\frac{(\frac{2k-n}{\sqrt{n}})^2}{2}} \frac{2}{\sqrt{n}}| + \frac{\epsilon}{2}.$$

By Stirling's Formula,

$$B_{n,k} := (\frac{1}{2})^n \binom{n}{k} = (\frac{1}{2})^n (1 + O(\frac{1}{n})) \frac{\sqrt{2\pi} n^{n+\frac{1}{2}}}{(2\pi) k^{k+\frac{1}{2}} (n-k)^{(n-k)+\frac{1}{2}}}$$

Letting $k = \frac{n}{2} + j$,

$$\begin{aligned} B_{n,k} &= (\frac{1}{2})^n (1 + O(\frac{1}{n})) \frac{n^{n+\frac{1}{2}}}{\sqrt{2\pi} (\frac{n}{2} + j)^{\frac{n}{2}+j+\frac{1}{2}} (\frac{n}{2} - j)^{\frac{n}{2}-j+\frac{1}{2}}} \\ &= (\frac{1}{2})^n (1 + O(\frac{1}{n})) \frac{n^{n+\frac{1}{2}}}{\sqrt{2\pi} (\frac{n}{2})^{n+1} (1 + \frac{2j}{n})^{\frac{n}{2}+j+\frac{1}{2}} (1 - \frac{2j}{n})^{\frac{n}{2}-j+\frac{1}{2}}} \\ &= (1 + O(\frac{1}{n})) \frac{2}{\sqrt{2\pi} \sqrt{n} (1 + \frac{2j}{n})^{\frac{n}{2}+j+\frac{1}{2}} (1 - \frac{2j}{n})^{\frac{n}{2}-j+\frac{1}{2}}}. \end{aligned}$$

It follows that

$$
\begin{aligned}
(1+\frac{2j}{n})^{\frac{n}{2}+j}(1-\frac{2j}{n})^{\frac{n}{2}-j} &= e^{(\frac{n}{2}+j)ln(1+\frac{2j}{n})}e^{(\frac{n}{2}-j)ln(1-\frac{2j}{n})} \\
&= e^{(\frac{n}{2}+j)(\frac{2j}{n}-\frac{2j^2}{n^2}+O(n^{\frac{3}{2}}))}e^{(\frac{n}{2}-j)(\frac{-2j}{n}-\frac{2j^2}{n^2}-O(n^{\frac{3}{2}}))} \\
&= e^{\frac{2j^2}{n}+O(\frac{1}{n})}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
(\frac{1}{2})^n\binom{n}{\frac{n}{2}+j} &= (1+O(\frac{1}{n}))\frac{2}{\sqrt{2\pi}\sqrt{n}}e^{\frac{-2j^2}{n}+O(\frac{1}{n})}(1+\frac{2j}{n})^{-\frac{1}{2}}(1-\frac{2j}{n})^{-\frac{1}{2}} \\
&= (1+O(\frac{1}{n}))\frac{2}{\sqrt{2\pi}\sqrt{n}}e^{\frac{-2j^2}{n}+O(\frac{1}{n})}(1+\frac{j}{n}+O(\frac{1}{n}))(1-\frac{j}{n}+O(\frac{1}{n})) \\
&= (1+O(\frac{1}{n}))\frac{2}{\sqrt{2\pi}\sqrt{n}}e^{\frac{-2j^2}{n}+O(\frac{1}{n})}(1-\frac{j^2}{n^2}+O(\frac{1}{n})).
\end{aligned}
$$

Thus,

$$
\sum_{k=\frac{n}{2}-\frac{b\sqrt{n}}{2}}^{\frac{n}{2}+\frac{b\sqrt{n}}{2}}|(\frac{1}{2})^n\binom{n}{k}-\frac{1}{\sqrt{2\pi}}e^{-\frac{(\frac{2k-n}{\sqrt{n}})^2}{2}}\frac{2}{\sqrt{n}}|
$$

$$
=\sum_{j=-\frac{b\sqrt{n}}{2}}^{\frac{b\sqrt{n}}{2}}|(\frac{1}{2})^n\binom{n}{\frac{n}{2}+j}-\frac{1}{\sqrt{2\pi}}e^{-\frac{2j^2}{n}}\frac{2}{\sqrt{n}}|
$$

$$
=\frac{2}{\sqrt{n}\sqrt{2\pi}}\sum_{j=-\frac{b\sqrt{n}}{2}}^{\frac{b\sqrt{n}}{2}}e^{-\frac{2j^2}{n}}|(1+O(\frac{1}{n}))^2(1-\frac{j^2}{n^2}+O(\frac{1}{n}))-1|
$$

$$
\leq\frac{4}{\sqrt{n}\sqrt{2\pi}}\sum_{j=0}^{\frac{b\sqrt{n}}{2}}(\frac{j^2}{n^2}+O(\frac{1}{n}))
$$

$$
=\frac{4}{\sqrt{n}\sqrt{2\pi}}(\frac{\frac{b\sqrt{n}}{2}(\frac{b\sqrt{n}}{2}+1)(b\sqrt{n}+1)}{6}+O(\frac{1}{\sqrt{n}}))
$$

$$
=\frac{b^3}{6\sqrt{2\pi}}\frac{1}{n}+O(\frac{1}{n}).
$$

Choosing n sufficiently large, we have $\frac{b^3}{6\sqrt{2\pi}}\frac{1}{n} < \frac{\epsilon}{2\|f\|_\infty}$. That is, we have proven the following theorem:
For each $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$n \geq N \Rightarrow |\sum_{k=0}^{n}(\frac{1}{2})^n\binom{n}{k}f(\frac{2k-n}{\sqrt{n}}) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{\frac{-x^2}{2}}dx| < \epsilon + O(\frac{1}{n}).$$

In the next section, we extend this proof to the general CLT by expanding our random variables using the Haar wavelet basis $\{H_{j,k}(x)|0 \leq j < \infty, 0 \leq k \leq 2^j - 1\} \cup \{\chi_{[0,1]}\}$ defined in the next section. For fixed $j \in \mathbb{N} \cup \{0\}$, we have

$$2^{-\frac{j}{2}}\sum_{k=0}^{2^j-1} H_{j,k}(x) = \epsilon_j$$

with $P(\epsilon_j = 1) = P(\epsilon_j = -1) = \frac{1}{2}$. Thus, the Haar wavelet basis is implicitlyembedded in this proof for Bernoulli random variables.

1.5 Creating an i.i.d. sequence of random variables on the probability space $([0,1], \mathcal{B}, \lambda)$

Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$. In this section, we will create an i.i.d. sequence of random variables on $([0,1], \mathcal{B}, \lambda)$, having the same distribution as X. Here, $\mathcal{B}$ denotes the sigma algebra of Borel subsets of $[0,1]$ and λ denotes Lebesque measure on $[0,1]$. Consider the cumulative distribution function $F : \mathbb{R} \to [0,1]$ given by $F(x) = P(X \leq x)$. We define $Q : [0,1] \to \mathbb{R}$ so that $Q(p) = \inf\{x \in \mathbb{R}|P(X \leq x) \geq p\}$. Note that Q is not a true inverse for F, as F is not injective. Q is often refered to as the quantile function for X. We can see that Q is a random variable on $([0,1], \mathcal{B}, \lambda)$ with the same distribution as X.

From X, we create an i.i.d. sequence of random variables by appropriately randomizing the binary sequence. Consider the binary expansion of $x \in [0,1]$ given by $(\epsilon(x)_i)_{i=1}^{\infty}$. We then have

$$x = \sum_{i=1}^{\infty}\frac{\epsilon_i(x)}{2^i}.$$

For an example, consult Appendix D.

We now create the following arrangement of $(\epsilon_i(x))_{i=1}^{\infty}$.

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 & \epsilon_3 \\ \epsilon_4 & \epsilon_5 & \epsilon_6 & \epsilon_7 \\ \epsilon_8 & \epsilon_9 & \epsilon_{10} & \epsilon_{11} & \epsilon_{12} & \epsilon_{13} & \epsilon_{14} & \epsilon_{15} \end{pmatrix}$$

One can see that row $i \in \mathbb{N}$ has 2^i elements numbered from left to right, starting with $\epsilon_{2^{i-1}}$. Any arrangement of (ϵ_i) without repetitions will work for our problem.

Now, we randomize $[0, 1]$ by forming the functions $P_i(x) : [0, 1] \to [0, 1]$, $i \in \mathbb{N}$, as follows: For each $i \in \mathbb{N}$ and $x \in [0, 1]$, let $P_i(x)$ denote the number in $[0, 1]$ whose binary expansion is obtained from the elements of column i, starting with the first entry. For example, the number $(P_2(x))$ would have the binary expansion $(\epsilon_3(x), \epsilon_5(x), \epsilon_9(x), ...)$. Since $\frac{3}{8}$ has the binary expansion $(0, 1, 1, 0, 0, 0, ...)$, then $P_2(\frac{3}{8})$ has the binary expansion $(1, 0, 0, 0, ...)$. Therefore, $P_2(\frac{3}{8}) = \frac{1}{2}$. Observe that $\lambda(\epsilon_j(P_i(x)) = 0) = \lambda(\epsilon_j(P_i(x)) = 1) = \frac{1}{2}$ for each $i, j \in \mathbb{N}$. Furthermore, the sequence $(P_i(x))$ is independent as each $P_i(x)$ depends on a disjoint set of ϵ_j's. We have thus created an i.i.d. sequence of random variables on [0,1] given by $X_i(x) = X(P_i(x))$ for each $i \in \mathbb{N}$.

1.6 Properties of the Haar basis

The Haar basis is the simplest orthonormal system on $[0, 1]$ and consists of the set $S = \{H_{j,k}(x) | 0 \le j < \infty, 0 \le k \le 2^j - 1\} \cup \{\chi_{[0,1]}\}$ where

$$H_{j,k} = \begin{cases} 2^{\frac{j}{2}} & x \in [\frac{k}{2^j}, \frac{k+\frac{1}{2}}{2^j}) \\ -2^{\frac{j}{2}} & x \in [\frac{k+\frac{1}{2}}{2^j}, \frac{k+1}{2^j}) \\ 0 & \text{otherwise} \end{cases}$$

As $S = \{H_{j,k}(x) | 0 \le j < \infty, 0 \le k \le 2^j - 1\} \cup \{\chi_{[0,1]}\}$ forms a complete orthonormal basis for $L^2([0, 1])$, we can expand any $\phi \in L^2([0, 1])$ as

$$\phi(x) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k} H_{j,k}(x) + \int_0^1 \phi(x)dx$$

where $c_{j,k} = \int_0^1 \phi(x) H_{j,k}(x)dx$. Let X be a random variable (measurable function) defined on the probability space $([0,1], \mathcal{B}, \lambda)$. If $X \in L^1([0,1], \mathcal{B}, \lambda)$, we may assume $EX = 0$ by simply changing X into $X - EX$. Assuming that $EX^2 < \infty$, $\text{var}X = \|X\|_2^2 < \infty$. Thus, we have $\int_0^1 X(x)dx = EX = 0$ and

$$X(x) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k} H_{j,k}(x)$$

where $c_{j,k} = \int_0^1 X(x) H_{j,k}(x)dx$.

For simplicity, let us assume that $\text{var}X = 1$ as we may replace X by $\frac{X}{\sqrt{\text{var}X}}$. Then, by Plancherel's equality, we have

$$\sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k}^2 = 1.$$

The Haar basis is discussed in more detail in [8]. Expanding $X(P_i(x))$ using the Haar basis, we have

$$X_i(x) = X(P_i(x)) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k} H_{j,k}(P_i(x)).$$

As $H_{j,k}(x) = 2^{\frac{j}{2}}(-1)^{\epsilon_{j+1}(x)} \chi_{\{k\}}(\lfloor 2^j x \rfloor)$,

$$X(P_i(x)) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k} 2^{\frac{j}{2}} (-1)^{\epsilon_{j+1}(P_i(x))} \chi_{\{k\}}(\lfloor 2^j P_i(x) \rfloor).$$

Setting $k = \lfloor 2^j P_i(x) \rfloor$,

$$X(P_i(x)) = \sum_{j=0}^{\infty} c_{j, \lfloor 2^j P_i(x) \rfloor} 2^{\frac{j}{2}} (-1)^{\epsilon_{j+1}(P_i(x))}.$$

1.7 Properties of $Y_M(x)$

The next two lemmas will investigate the properties of the random variable

$$Y_M(x) = \sum_{j=0}^{M} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x \rfloor} (-1)^{\epsilon_{j+1}(x)}$$

for $x \in [0, 1]$. The random variable Y_M plays an essential role in proving the CLT. Y_M is a discrete random variable with finitely many values. Note that this function depends on $(\epsilon_1(x), ..., \epsilon_{M+1}(x))$. Therefore, $Y_M(x)$ is constant on dyadic intervals of the form $[\frac{k}{2^{M+1}}, \frac{k+1}{2^{M+1}})$ for $0 \leq k \leq 2^{M+1} - 1$. We call the 2^{M+1} values of the function $Y_M(x)$ outcomes. For an example, consider the case where $M = 2$. We will then find the possible values for the function

$$Y_2(x) = \sum_{j=0}^{2} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x \rfloor} (-1)^{\epsilon_{j+1}(x)}.$$

Since $Y_2(x)$ depends on $(\epsilon_1(x), \epsilon_2(x), \epsilon_3(x))$, then $Y_2(x)$ is constant on the dyadic intervals $[\frac{k}{8}, \frac{k+1}{8})$ for $0 \leq k \leq 7$. We have the following outcomes for $Y_2(x)$:

$$\begin{aligned}
[0, \frac{1}{8}) : o_1 &= c_{0,0} + \sqrt{2}c_{1,0} + 2c_{2,0} \\
[\frac{1}{8}, \frac{2}{8}) : o_2 &= c_{0,0} + \sqrt{2}c_{1,0} - 2c_{2,0} \\
[\frac{2}{8}, \frac{3}{8}) : o_3 &= c_{0,0} - \sqrt{2}c_{1,0} + 2c_{2,1} \\
[\frac{3}{8}, \frac{4}{8}) : o_4 &= c_{0,0} - \sqrt{2}c_{1,0} - 2c_{2,1} \\
[\frac{4}{8}, \frac{5}{8}) : o_5 &= -c_{0,0} + \sqrt{2}c_{1,1} + 2c_{2,2} \\
[\frac{5}{8}, \frac{6}{8}) : o_6 &= -c_{0,0} + \sqrt{2}c_{1,1} - 2c_{2,2} \\
[\frac{6}{8}, \frac{7}{8}) : o_7 &= -c_{0,0} - \sqrt{2}c_{1,1} + 2c_{2,3} \\
[\frac{7}{8}, 1) : o_8 &= -c_{0,0} - \sqrt{2}c_{1,1} + 2c_{2,3}
\end{aligned}$$

Observe that $\sum_{i=1}^{8} o_i = 0$.

Lemma 1.7.1. *If $(o_i)_{i=1}^{2^{M+1}}$ are outcomes of Y_M, then $\sum_{i=1}^{2^{M+1}} o_i = 0$.*

Proof. For $x \in [0,1]$, we have $n = 2^{M+1}$ possible outcomes, each corresponding to an interval of the form $[\frac{k}{2^{M+1}}, \frac{k+1}{2^{M+1}})$ for some $0 \le k \le 2^{M+1}-1$. Each outcome is equal to a sum of $M+1$ terms. Now, we fix $j = j_0$ and consider this term in the outcome. For $j = j_0$, we consider terms of the form

$$2^{\frac{j_0}{2}} c_{j_0,\lfloor 2^{j_0}x\rfloor}(-1)^{\epsilon_{j_0+1}(x)}.$$

We note the $c_{j_0,\lfloor 2^{j_0}x\rfloor}$ depends only on $(\epsilon_1(x), \ldots \epsilon_{j_0}(x))$ as $\lfloor 2^{j_0}x\rfloor = k$ if and only if $x \in [\frac{k}{2^{j_0}}, \frac{k+1}{2^{j_0}})$. Thus, $c_{j_0,\lfloor 2^{j_0}x\rfloor}$ is constant on intervals of the form $[\frac{k}{2^{j_0}}, \frac{k+1}{2^{j_0}})$ for $0 \le k \le 2^{j_0}-1$. Each interval $[\frac{k}{2^{j_0}}, \frac{k+1}{2^{j_0}})$ is the union of 2^{M+1-j_0} intervals of length $\frac{1}{2^{M+1}}$, and hence the coefficient $c_{j_0,k}$ corresponds to 2^{M+1-j_0} outcomes. If we cut the interval $[\frac{k}{2^{j_0}}, \frac{k+1}{2^{j_0}})$ in half, then for the left half we have $\epsilon_{j_0+1}(x) = 0$ and for the right half we have $\epsilon_{j_0+1}(x) = 1$. Thus, we have 2^{M-j_0} outcomes with the coefficient $2^{\frac{j_0}{2}} c_{j_0,k}$ and 2^{M-j_0} outcomes with the coefficient $-2^{\frac{j_0}{2}} c_{j_0,k}$ for each $0 \le k \le 2^{j_0}-1$. Summing over all $0 \le k \le 2^{j_0}-1$ and then over all $0 \le j_0 \le M$, we have $\sum_{i=1}^{n} o_i = 0$. □

Lemma 1.7.2. *If $(o_i)_{i=1}^{2^{M+1}}$ are outcomes of Y_M, then $\sum_{i=1}^{2^{M+1}} o_i^2 = 2^{M+1}\sigma_M^2$, where*

$$\sigma_M = \sqrt{\sum_{j=0}^{M}\sum_{k=0}^{2^j-1} c_{j,k}^2}.$$

Proof. Group the outcomes in pairs, corresponding to the dyadic intervals $[\frac{k}{2^{M+1}}, \frac{k+1}{2^{M+1}})$ and $[\frac{k+1}{2^{M+1}}, \frac{k+2}{2^{M+1}})$ for k even. Furthermore, for each outcome, group the first M terms together. Therefore, for $x_1 \in [\frac{k}{2^{M+1}}, \frac{k+1}{2^{M+1}})$, we have

$$\left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_1\rfloor}(-1)^{\epsilon_{j+1}(x_1)} + 2^{\frac{M}{2}} c_{M,\lfloor 2^M x_1\rfloor}\right),$$

and for $x_2 \in [\frac{k+1}{2^{M+1}}, \frac{k+2}{2^{M+1}})$ we have

$$\left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_2\rfloor}(-1)^{\epsilon_{j+1}(x_2)} - 2^{\frac{M}{2}} c_{M,\lfloor 2^M x_2\rfloor}\right).$$

Squaring and then adding these terms, we obtain

$$\left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_1\rfloor}(-1)^{\epsilon_{j+1}(x_1)}\right)^2 + 2\left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_1\rfloor}(-1)^{\epsilon_{j+1}(x_1)}\right)\left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_1\rfloor}\right)$$

$$+\left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_1\rfloor}\right)^2 + \left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_2\rfloor}(-1)^{\epsilon_{j+1}(x_2)}\right)^2$$

$$-2\left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_2\rfloor}(-1)^{\epsilon_{j+1}(x_2)}\right)\left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_2\rfloor}\right) + \left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_2\rfloor}\right)^2.$$

Since $x_1, x_2 \in [\frac{l}{2^M}, \frac{l+1}{2^M})$ for $2l = k$, then we have $\lfloor 2^j x_1\rfloor = \lfloor 2^j x_2\rfloor$ and $\epsilon_j(x_1) = \epsilon_j(x_2)$ for all $0 \le j \le M$. Therefore, we have

$$2\left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_1\rfloor}(-1)^{\epsilon_{j+1}(x_1)}\right)\left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_1\rfloor}\right)$$

$$-2\left(\sum_{j=0}^{M-1} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_2\rfloor}(-1)^{\epsilon_{j+1}(x_2)}\right)\left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_2\rfloor}\right) = 0.$$

As $\lfloor 2^M x_1\rfloor = \lfloor 2^M x_2\rfloor = l$, summing the squares of the two final terms, we have

$$\left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_1\rfloor}\right)^2 + \left(2^{\frac{M}{2}} c_{M,\lfloor 2^M x_2\rfloor}\right)^2$$

$$= 2^M c_{M,l}^2 + 2^M c_{M,l}^2 = 2^{M+1} c_{M,l}^2.$$

Summing over all terms corresponding to $j = M$, we have

$$2^{M+1}\left(\sum_{l=0}^{2^M-1} c_{M,l}^2\right).$$

Now, we consider

$$\sum_{i=1}^{2}\left(\sum_{j=0}^{M-2} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_i\rfloor}(-1)^{\epsilon_{j+1}(x_i)} + 2^{\frac{M-1}{2}} c_{M-1,\lfloor 2^{M-1} x_i\rfloor}\right)^2$$

for some $x_1, x_2 \in [\frac{k}{2^M}, \frac{k+1}{2^M})$ for even k. Then, we choose $x_3, x_4 \in [\frac{2k+1}{2^M}, \frac{2k+2}{2^M})$ and consider

$$\sum_{i=3}^{4} \left(\sum_{j=0}^{M-2} 2^{\frac{j}{2}} c_{j,\lfloor 2^j x_i \rfloor} (-1)^{\epsilon_{j+1}(x_i)} + 2^{\frac{M-1}{2}} c_{M-1,\lfloor 2^{M-1} x_i \rfloor} \right)^2 .$$

Adding these terms together, we now obtain

$$\left(2^{\frac{M-1}{2}} c_{M-1,\lfloor 2^{M-1} x_1 \rfloor}\right)^2 + \left(2^{\frac{M-1}{2}} c_{M-1,\lfloor 2^{M-1} x_2 \rfloor}\right)^2$$

$$+ \left(2^{\frac{M-1}{2}} c_{M-1,\lfloor 2^{M-1} x_3 \rfloor}\right)^2 + \left(2^{\frac{M-1}{2}} c_{M-1,\lfloor 2^{M-1} x_4 \rfloor}\right)^2$$

$$= 2^{M-1} c_{M-1,l}^2 + 2^{M-1} c_{M-1,l}^2 + 2^{M-1} c_{M-1,l}^2 + 2^{M-1} c_{M-1,l}^2$$

$$= 4 \cdot 2^{M-1} c_{M-1,l}^2 = 2^{M+1} c_{M-1,l}^2.$$

It follows that

$$\sum_{i=1}^{n} o_i^2 = 2^{M+1} \left(\sum_{j=0}^{M} \sum_{k=0}^{2^j-1} c_{j,k}^2 \right) = 2^{M+1} \sigma_M^2.$$

□

1.8 A new proof of the CLT

In this section, we will give a new proof of the Central Limit Theorem:

Theorem 1.8.1. *Let (X_i) be a sequence of i.i.d. random variables with mean μ and variance σ^2. Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded, continuous function. Then, for each $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that*

$$\left| E\left(f\left(\frac{X_1 + ... + X_N - N\mu}{\sigma\sqrt{N}} \right) \right) - E(f(Y)) \right| < \frac{\|f\|_\infty C n^{\frac{5}{2}} b^{n+1}}{4(2\pi)^{\frac{n-1}{2}} N},$$

where $Y \overset{d}{=} N(0,1)$, for all $N \geq N_0$, where b and n will be defined later.

Proof. The proof proceeds via the following steps, A-G:

- **A.** We truncate the Haar expansions for our random variables to have only finitely many terms.
- **B.** We examine the truncated Haar expansion and show that it is actually a multinomial random variable.
- **C.** We cut off the tails of the multinomial random variable by using Chebyshev's inequality.
- **D.** We use Stirling's formula and Taylor series to approximate the multinomial coefficients.
- **E.** We write our Gaussian random variable as a sum of independent Gaussian random variables and then express the expected value as an integral.
- **F.** We cut off the tails of our Gaussian integral and then express this integral as a Riemann sum.
- **G.** We compute the difference between the expected value for the sums of our truncated Haar expansions and the Gaussian Riemann sum.

Step A. Let

$$y(N) := \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sum_{j=0}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x) \rfloor} (-1)^{\epsilon_{j+1}(P_i(x))},$$

$$x(N, M) := \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sum_{j=0}^{M} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x) \rfloor} (-1)^{\epsilon_{j+1}(P_i(x))},$$

$$z(N, M) := \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x) \rfloor} (-1)^{\epsilon_{j+1}(P_i(x))}.$$

Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded, continuous function. Let $\epsilon > 0$. We have

$$E(y(N)) = 0.$$

By Chebyshev's inequality, we have

$$P\left(|y(N)| > L\right) = \frac{\text{var}(y(N))}{L^2} = \frac{1}{L^2} < \epsilon$$

for sufficiently large L. Note that $P(|y(N)| > L) < \epsilon$ for all $N \in \mathbb{N}$. Therefore, by uniform continuity of f on $[-L, L]$, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $|x|,|y| \leq L$ and $|x - y| < \delta$.

We will now compute

$$\text{var}(z(N,M)) = \text{var}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x)\rfloor}(-1)^{\epsilon_{j+1}(P_i(x))}\right)$$

$$= \frac{1}{N}\text{var}\left(\sum_{i=1}^{N}\sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x)\rfloor}(-1)^{\epsilon_{j+1}(P_i(x))}\right).$$

As

$$\left(\sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x)\rfloor}(-1)^{\epsilon_{j+1}(P_i(x))}\right)$$

is an independent sequence,

$$\text{var}(z(N,M)) = \frac{1}{N}\sum_{i=1}^{N}\text{var}\left(\sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x)\rfloor}(-1)^{\epsilon_{j+1}(P_i(x))}\right).$$

As

$$\left(\sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_i(x)\rfloor}(-1)^{\epsilon_{j+1}(P_i(x))}\right)$$

is an identically distributed sequence,

$$\text{var}(z(N,M)) = \frac{1}{N}\cdot N\cdot\text{var}\left(\sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_1(x)\rfloor}(-1)^{\epsilon_{j+1}(P_1(x))}\right)$$

$$= \text{var}\left(\sum_{j=M+1}^{\infty} 2^{\frac{j}{2}} c_{j,\lfloor 2^j P_1(x)\rfloor}(-1)^{\epsilon_{j+1}(P_1(x))}\right)$$

$$= \text{var}\left(\sum_{j=M+1}^{\infty}\sum_{k=0}^{2^j-1} c_{j,k} H_{j,k}(P_1(x))\right).$$

By Plancherel's formula,

$$\mathrm{var}(z(N,M)) = \sum_{j=M+1}^{\infty} \sum_{k=0}^{2^j-1} c_{j,k}^2 = 1 - \sigma_M^2.$$

Note that $\mathrm{var}(z(N,M))$ does not depend on N.

By Cheyshev's Inequality, we have

$$P\left(|y(N) - x(N,M)| \geq \delta\right) = P\left(|z(N,M)| \geq \delta\right) \leq \frac{\mathrm{var}(z(N,M))}{\delta^2} = \frac{1-\sigma_M^2}{\delta^2} < \epsilon$$

for sufficiently large M.

Then, we have

$$P\left(|y(N) - x(N,M)| \geq \delta\right) \leq \epsilon.$$

Let

$$A_M = \{|y(N) - x(N,M)| \geq \delta\}.$$

Then, we have

$$|E_P(f(y(N))) - E_P(f(x(N,M)))|$$

$$\leq \left|E_P\left((f(y(N)) - f(x(N,M)))\chi_{A_M}\right)\right| + |E_P((f(y(N)) - f(x(N,M)))\chi_{A_M^C})|$$

$$\leq E_P|(f(y(N)) - f(x(N,M)))\chi_{A_M}| + E_P|(f(y(N)) - f(x(N,M)))\chi_{A_M^C}|$$

$$\leq 2\|f\|_\infty P(A_M) + \epsilon P(A_M^C)$$

$$\leq \epsilon(2\|f\|_\infty + 1).$$

If $X \stackrel{d}{=} N(0,1)$, then

$$\sigma_M X \to X$$

as $M \to \infty$. Thus, by continuity of f,

$$f(\sigma_M X) \to f(X)$$

as $M \to \infty$. Hence, by the Dominated Convergence Theorem,

$$E(f(\sigma_M X)) \to E(f(X))$$

as $M \to \infty$. Hence, there exists $M_0 \in \mathbb{N}$ such that

$$|E(f(\sigma_M X)) - E(f(X))| < \epsilon$$

for all $M \geq M_0$.

So let M be large enough so that

$$\frac{1 - \sigma_M^2}{\delta^2} < \epsilon$$

and

$$|E(f(\sigma_M X)) - E(f(X))| < \epsilon.$$

Note that M does not depend on N. M and n will be fixed from now on.

Step B. We will now take a closer look at $x(N, M)$. Let

$$x(N, M) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Y(M, i),$$

and note that

$$Y(M, i) = \sum_{j=0}^{M} 2^{\frac{j}{2}} c_{j, \lfloor 2^j P_i(x) \rfloor} (-1)^{\epsilon_{j+1}(P_i(x))}.$$

Then, $Y(M, i)$ is a random variable with $n = 2^{M+1}$ possible values $o_1, ..., o_n$ each having probability $\frac{1}{n}$. As $N \to \infty$, we can assume that $N > n$. Therefore, the outcomes of $Y(M, i)$ must repeat. If k_i denotes the number of the outcome o_i in the sum

$$\sum_{i=1}^{N} Y(M, i),$$

then this random variable has values

$$k_1 o_1 + ... + k_n o_n$$

where

$$k_1 + ... + k_n = N.$$

Therefore, for fixed $k_1, ..., k_n$ with $k_1 + ... + k_n = N$,

$$P\left(\sum_{i=1}^{N} Y(M, i) = \sum_{i=1}^{n} k_i o_i\right) = \binom{N}{k_1, ..., k_n} \left(\frac{1}{n}\right)^N,$$

and

$$E_P\left(f(x(N,M))\right) = \sum_{\substack{k_1=1 \\ k_1+\ldots+k_n=N}}^{N} \ldots \sum_{k_n=1}^{N} \frac{1}{n^N} \binom{N}{k_1, \ldots, k_n} f\left(\frac{\sum_{i=1}^{n} k_i o_i}{\sqrt{N}}\right).$$

Step C. We will now cut off the tails from our expected value using Chebyshev's inequality. Let K_i be the random variable which denotes the number of times the outcome o_i is observed, having values k_i. We have $E(K_i) = N(\frac{1}{n}) = \frac{N}{n}$ and $\text{var}(K_i) = N(\frac{1}{n})(1-\frac{1}{n})$. By Chebyshev's Inequality, we have

$$P\left(|K_i - \frac{N}{n}| \geq b_1\sqrt{\frac{N}{n}(1-\frac{1}{n})}\right) \leq \frac{\sqrt{\frac{N}{n}(1-\frac{1}{n})}}{b_1\sqrt{\frac{N}{n}(1-\frac{1}{n})}} = \frac{1}{b_1} < \frac{\epsilon}{n}$$

for large enough $b_1 > 0$. Let $b = \max\{b_0, b_1\}$, where b_0 will be defined in Step F. Set

$$A_{N,n} := \{|K_i - \frac{N}{n}| \geq b\sqrt{\frac{N}{n}(1-\frac{1}{n})} \text{ for some i, } 1 \leq i \leq n\}.$$

Then, by finite additivity, we have

$$P\left(A_{N,n}\right) < n\left(\frac{\epsilon}{n}\right) = \epsilon.$$

Let $h(N,n) := \sqrt{\frac{N}{n}(1-\frac{1}{n})}$. Let

$$q(N, k_1, \ldots k_n) := \frac{1}{n^N}\binom{N}{k_1, \ldots, k_n} f\left(\frac{\sum_{i=1}^{n} k_i o_i}{\sqrt{N}}\right).$$

Set

$$E_i := \{1 \leq K_i \leq N \text{ and } \sum_{i=1}^{n} K_i = N\}.$$

Set

$$B_{N,n} := A_{N,n} \cap \left(\cup_{i=1}^{n} E_i\right).$$

Then, we have

$$P(B_{N,n}) \leq P(A_{N,n}) < \epsilon.$$

It follows that

$$\left| \sum_{\substack{k_1=1 \\ k_1+\ldots+k_n=N}}^{N} \ldots \sum_{k_n=1}^{N} q(N, k_1, \ldots k_n) - \sum_{k_1=\frac{N}{n}-b_1h(N,n)}^{\frac{N}{n}+b_1h(N,n)} \ldots \sum_{\substack{k_n=\frac{N}{n}-b_1h(N,n) \\ k_1+\ldots+k_n=N}}^{\frac{N}{n}+b_1h(N,n)} q(N, k_1, \ldots, k_n) \right| < \epsilon \|f\|_\infty.$$

Note that

$$E(x(N, M)) = E\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} o_i K_i\right) =$$

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} o_i E(K_i) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} o_i \frac{N}{n} = \frac{1}{\sqrt{N}} \frac{N}{n} \sum_{i=1}^{n} o_i = 0.$$

and

$$\operatorname{var}(x(N, M)) = \operatorname{var}\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} o_i K_i\right) =$$

$$= \frac{1}{N} \sum_{i=1}^{N} o_i^2 \operatorname{var}(K_i) = \frac{1}{N}(N)\left(\frac{1}{n}\right)\left(1 - \frac{1}{n}\right) \sum_{i=1}^{n} o_i^2$$

$$= \left(\frac{1}{n}\right)\left(1 - \frac{1}{n}\right) n\sigma_M^2 = \left(1 - \frac{1}{n}\right) \sigma_M^2.$$

Step D. We will now use Stirling's formula and Taylor series to approximate the multinomial coefficients. Set

$$l(N, k_1, \ldots, k_n) := \frac{1}{n^N} \binom{N}{k_1, \ldots, k_n}.$$

By Stirling's Formula, we have

$$l(N, k_1, \ldots, k_n) = \frac{\left(1 + O\left(\frac{1}{N}\right)\right)(2\pi)^{\frac{1}{2}} N^{N+\frac{1}{2}}}{(2\pi)^{\frac{n}{2}} n^N (N + k_1)^{(N+k_1+\frac{1}{2})} + \ldots + (N + k_n)^{(N+k_n+\frac{1}{2})}}.$$

Letting $k_i = \frac{N}{n} + j_i$ for $1 \le i \le n$,

$$l(N, k_1, \ldots, k_n) = \frac{\left(1 + O\left(\frac{1}{N}\right)\right)(2\pi)^{\frac{1}{2}} N^{N+\frac{1}{2}}}{(2\pi)^{\frac{n}{2}} n^N (\frac{N}{n} + j_1)^{(\frac{N}{n}+j_1+\frac{1}{2})} + \ldots + (\frac{N}{n} + j_n)^{(\frac{N}{n}+j_n+\frac{1}{2})}}.$$

Factoring $(\frac{N}{n})$ from each term $(\frac{N}{n}+j_i)$ in the denominator,

$$l(N,k_1,...,k_n) = \frac{\left(1+O\left(\frac{1}{N}\right)\right) N^{N+\frac{1}{2}}}{(2\pi)^{\frac{n-1}{2}} n^N (\frac{N}{n})^{N+\frac{n}{2}} (1+\frac{nj_1}{N})^{(\frac{N}{n}+j_1+\frac{1}{2})} + ... + (1+\frac{nj_n}{N})^{(\frac{N}{n}+j_n+\frac{1}{2})}}$$

$$= \frac{\left(1+O\left(\frac{1}{N}\right)\right) n^{\frac{n}{2}}}{(2\pi)^{\frac{n-1}{2}} N^{\frac{n-1}{2}} (1+\frac{nj_1}{N})^{(\frac{N}{n}+j_1+\frac{1}{2})} + ... + (1+\frac{nj_n}{N})^{(\frac{N}{n}+j_n+\frac{1}{2})}}.$$

For all $1 \le i \le n$, we set

$$m(N,n,i) := (1+\frac{nj_i}{N})^{\frac{N}{n}+j_i+\frac{1}{2}} = e^{(\frac{N}{n}+j_i+\frac{1}{2}) ln(1+\frac{nj_i}{N})}.$$

Using a Taylor series approximation,

$$\begin{aligned} m(N,n,i) &= e^{(\frac{N}{n}+j_i+\frac{1}{2})(\frac{nj_i}{N}-\frac{n^2 j_i^2}{2N^2}+O(\frac{1}{N^3}))} \\ &= e^{j_i+\frac{j_i^2 n}{2N}+\frac{nj_i}{2N}-\frac{n^2 j_i^3}{2N^2}-\frac{n^2 j_i^2}{4N^2}+O(\frac{1}{N^2})}. \end{aligned}$$

Therefore, we have

$$(1+\frac{nj_1}{N})^{\frac{N}{n}+j_1+\frac{1}{2}} ... (1+\frac{nj_n}{N})^{\frac{N}{n}+j_n+\frac{1}{2}} = e^{(\frac{n}{2N}-\frac{n^2}{4N^2})\sum_{\alpha=1}^n j_\alpha^2 - \frac{n^2}{2N^2}\sum_{\alpha=1}^n j_\alpha^3 + O(\frac{1}{N^2})}.$$

Set

$$H(N,j_1,...,j_n) := (-\frac{n}{2N}+\frac{n^2}{4N^2})\sum_{i=1}^n j_i^2 + \frac{n^2}{2N^2}\sum_{i=1}^n j_i^3 + O(\frac{1}{N^2}).$$

Hence, we have

$$l(N,j_1,...,j_n) = \left(1+O\left(\frac{1}{N}\right)\right) \frac{e^{H(N,j_1,...,j_n)} n^{\frac{n}{2}}}{(2\pi)^{\frac{n-1}{2}} N^{\frac{n-1}{2}}}.$$

Step E. Let $Y_1,...,Y_n$ be Gaussian random variables with mean 0 and variance 1. Let $(o_i)_{i=1}^n$ be as in Section 1.7. Then, consider $o_i' = \frac{o_i}{\sigma_M}$. For simplicity of notation, we will replace o_i' by o_i. Then, we have

$$\sum_{i=1}^n o_i^2 = n, \quad \sum_{i=1}^n o_i = 0.$$

Therefore, we have $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} o_i Y_i \overset{d}{=} N(0,1)$.

Set

$$O = \begin{bmatrix} o_1 \\ . \\ . \\ o_n \end{bmatrix}, \quad Y = \begin{bmatrix} Y_1 \\ . \\ . \\ Y_n \end{bmatrix}.$$

Then, the vector projection of Y in direction of the vector

$$u = \begin{bmatrix} 1 \\ . \\ . \\ 1 \end{bmatrix}$$

is given by

$$V = \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) u = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i\right)\frac{u}{\sqrt{n}} = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i\right)\frac{u}{\|u\|}$$

which means that V is a one dimensional standard normal on the line through the origin and orthogonal to the hyperplane in $\mathbb{R}^n$ given by

$$\sum_{i=1}^{n} y_i = 0.$$

Call that hyperplane S. Viewing V and $Z := Y - V$ as vector valued random variables, in addition to being orthogonal, as vectors, they are also independent as random variables. This can be verified by checking that all components (or coordinates) of Z are independent of all components (coordinates) of V. Since all components of V are equal to

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

that is V is a one dimensional random variable, and components (coordinates) of $Y - V$ are $Y_i - \bar{Y}$, the independence of Gaussian random variables follows from

$$E((Y_i - \bar{Y})\bar{Y}) = E(Y_i\bar{Y}) - E\left(\bar{Y}\right)^2 = \frac{1}{n} - \frac{1}{n} = 0$$

for all $i = 1, ..., n$. Note that $O^T Y = O^T (Y - V) + O^T V$, and since $O^T u = 0$,

$$f\left(\frac{\sum_{i=1}^n o_i Y_i}{\sqrt{n}}\right) = f\left(\frac{1}{\sqrt{n}} O^T Y\right) = f\left(\frac{1}{\sqrt{n}} O^T Z\right),$$

that is

$$f\left(\frac{1}{\sqrt{n}} O^T Z\right)$$

does not depend on V. Note that since the components of Z satisfy

$$\sum_{i=1}^n \left(Y_i - \bar{Y}\right) = 0,$$

the law of Z, $\mathcal{L}(Z)$, is a standard Gaussian $n-1$ dimensional measure on the hyperplane S. From the independence of Z and V, and Fubini's Theorem, it follows that

$$Ef\left(\frac{1}{\sqrt{n}} O^T Y\right) = E_{\mathcal{L}(Z)} E_{\mathcal{L}(\mathcal{V})} f\left(\frac{1}{\sqrt{n}} O^T Z\right) = E_{\mathcal{L}(Z)} f\left(\frac{1}{\sqrt{n}} O^T Z\right),$$

Since the density of Y is given by

$$\frac{1}{\left(\sqrt{2\pi}\right)^n} \exp(-\frac{1}{2} y^T y) = \frac{1}{\left(\sqrt{2\pi}\right)^{n-1}} \exp(-\frac{1}{2} z^T z) \frac{1}{\left(\sqrt{2\pi}\right)} \exp\left(-\frac{1}{2} (\bar{y})^2\right),$$

where y, z, and $\bar{y}$ are the realizations of Y, Z, and $\bar{Y}$, respectively, it follows that

$$E_{\mathcal{L}(Z)} f\left(\frac{1}{\sqrt{n}} O^T Z\right) = \frac{1}{\left(\sqrt{2\pi}\right)^{n-1}} \int_S f\left(\frac{1}{\sqrt{n}} \sum_{n=1}^n o_i y_i\right) \exp\left(-\frac{1}{2} \sum_{i=1}^n y_i^2\right) dS,$$

i.e., the expected value with respect to $\mathcal{L}(Z)$ is a surface integral over the hyperplane S. In arriving at the last equality we have used that

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = 0 \text{ on } S.$$

By projecting S onto $y_n = 0$ plane we have

$$E_{\mathcal{L}(Z)} f\left(\frac{1}{\sqrt{n}} O^T Z\right) = \frac{\sqrt{n}}{\left(\sqrt{2\pi}\right)^{n-1}} \int \ldots \int f\left(\frac{1}{\sqrt{n}} \sum_{n=1}^n o_i y_i\right) e^{\left(-\frac{1}{2} \sum_{i=1}^n y_i^2\right)} dy_1 \ldots dy_{n-1}$$

where

$$y_n = -\sum_{i=1}^{n-1} y_i.$$

The factor $\sqrt{n}$ appears as the result of replacing dS by $dy_1...dy_{n-1}$.

Step F. We will now cut off the tails of our Gaussian integral and then approximate the integral by a Riemann sum. Choose $b_0 > 0$ such that

$$\sqrt{n}\left|\int ... \int \frac{1}{(\sqrt{2\pi})^{n-1}} e^{-\frac{1}{2}\left(\sum_{i=1}^{n} y_i^2\right)} dy_1...dy_{n-1} - \int_{-b_2}^{b_2} ... \int_{-b_2}^{b_2} \frac{1}{(\sqrt{2\pi})^{n-1}} e^{-\frac{1}{2}\left(\sum_{i=1}^{n} y_i^2\right)} dy_1...dy_{n-1}\right| < \epsilon$$

for all $b_2 \geq b_0$. Then, we consider

$$I = \frac{\sqrt{n}}{\left(\sqrt{2\pi}\right)^{n-1}} \int ... \int f\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} o_i y_i\right) e^{\left(-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right)} dy_1...dy_{n-1}.$$

and

$$I_b = \frac{\sqrt{n}}{\left(\sqrt{2\pi}\right)^{n-1}} \int_{-b}^{b} ... \int_{-b}^{b} f\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} o_i y_i\right) e^{\left(-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right)} dy_1...dy_{n-1}.$$

Let $g(y_1, ...y_n) := f\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} o_i y_i\right) e^{\left(-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right)}$. Note that

$$\left|\sum_{i=1}^{n} o_i y_i\right| \leq \sum_{i=1}^{n} |o_i||y_i|$$

$$\leq b \sum_{i=1}^{n} |o_i| < \infty.$$

Then, let

$$S = \frac{\sqrt{n}}{\left(\sqrt{2\pi}\right)^{n-1}} \sum_{j_1=-bh(N,n)}^{bh(N,n)} ... \sum_{\substack{j_{n-1}=-bh(N,n) \\ j_1+j_2+...+j_n=0}}^{bh(N,n)} \int_{j_1 \frac{\sqrt{n}}{\sqrt{N}}}^{(j_1+1)\frac{\sqrt{n}}{\sqrt{N}}} ... \int_{j_{n-1} \frac{\sqrt{n}}{\sqrt{N}}}^{(j_{n-1}+1)\frac{\sqrt{n}}{\sqrt{N}}} g(y_1, ..., y_n) dy_1...dy_{n-1}$$

where $h(N, n)$ is as in Step D. Let

$$B_{N,n} := \{\frac{1}{\sqrt{N}} \sum_{i=1}^{n} o_i j_i > \sum_{i=1}^{n} |o_i| b\}.$$

Then, we have

$$|I - \frac{n^{\frac{n}{2}}}{\left(\sqrt{2\pi}\right)^{n-1} N^{\frac{n-1}{2}}} \sum_{\substack{j_1=-bh(N,n) \\ j_1+j_2+...+j_n=0}}^{bh(N,n)} ... \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} f\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{n} o_i j_i\right) e^{\left(-\frac{n}{2}\sum_{i=1}^{n} \frac{j_i^2}{N}\right)}|$$

$$\leq |I - S| + |S - \frac{n^{\frac{n}{2}}}{\left(\sqrt{2\pi}\right)^{n-1} N^{\frac{n-1}{2}}} \sum_{\substack{j_1=-bh(N,n) \\ j_1+j_2+...+j_n=0}}^{bh(N,n)} ... \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} f\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{n} o_i j_i\right) e^{\left(-\frac{n}{2}\sum_{i=1}^{n} \frac{j_i^2}{N}\right)}|$$

$$< \epsilon + \epsilon P(B_{N,n}) + 2\|f\|_\infty P(B_{N,n}^C) \leq \epsilon(1 + 2\|f\|_\infty)$$

for large enough N.

Step G. Now, assuming that $\sum_{i=1}^{n} j_i = 0$, we consider the difference

$$D_N := |\sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} \frac{n^{\frac{n}{2}}}{(2\pi)^{\frac{n-1}{2}} N^{\frac{n-1}{2}}} e^{H(N,j_1,...,j_n)} f\left(\frac{\sum_{i=1}^{n} j_i o_i}{\sqrt{N}}\right)$$

$$- \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} \frac{n^{\frac{n}{2}}}{(2\pi)^{\frac{n-1}{2}} N^{\frac{n-1}{2}}} e^{\left(-\frac{n}{2}\sum_{i=1}^{n} \frac{j_i^2}{N}\right)} f\left(\frac{\sum_{i=1}^{n} j_i o_i}{\sqrt{N}}\right)|$$

$$= \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} |\sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} f\left(\sum_{i=1}^{n} \frac{j_i o_i}{\sqrt{N}}\right) \left(e^{H(N,j_1,...,j_n)} - e^{\left(-\frac{n}{2}\sum_{i=1}^{n} \frac{j_i^2}{N}\right)}\right)|$$

$$= \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} |\sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} f\left(\sum_{i=1}^{n} \frac{j_i o_i}{\sqrt{N}}\right) e^{\left(-\frac{n}{2}\sum_{i=1}^{n} \frac{j_i^2}{N}\right)} \left(e^{G(N,j_1,...,j_n)} - 1\right)|$$

where $G(N, j_1, ..., j_n) = (\frac{n^2}{4N^2}) \sum_{i=1}^{n} j_i^2 + \frac{n^2}{2N^2} \sum_{i=1}^{n} j_i^3 + O(\frac{1}{N^2})$. Then,

$$D_N \leq \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} |\sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} f\left(\sum_{i=1}^{n} \frac{j_i o_i}{\sqrt{N}}\right) \left(e^{G(N,j_1,...,j_n)} - 1\right)|$$

$$\leq \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} |f\left(\sum_{i=1}^{n} \frac{j_i o_i}{\sqrt{N}}\right) \left(e^{G(N,j_1,...,j_n)} - 1\right)|.$$

Since $G(N, j_1, ...j_n) \geq 0$, then $\left(e^{G(N,j_1,...,j_n)} - 1\right) \geq 0$ and thus

$$D_N \leq \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} |f\left(\sum_{i=1}^{n} \frac{j_i o_i}{\sqrt{N}}\right)| \left(e^{G(N,j_1,...,j_n)} - 1\right).$$

Since f is bounded,

$$D_N \leq \|f\|_\infty \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} \left(e^{G(N,j_1,...,j_n)} - 1\right).$$

Using a Taylor series approximation, we have

$$D_N \leq \|f\|_\infty \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-b(N,n)}^{bh(N,n)} \left(1 + G(N, j_1, ..., j_n) + O(\frac{1}{N^2}) - 1\right)$$

$$= \|f\|_\infty \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-b(N,n)}^{bh(N,n)} \left(G(N, j_1, ..., j_n) + O(\frac{1}{N^2})\right)$$

$$= \|f\|_\infty \frac{n^{\frac{n}{2}}}{(2\pi N)^{\frac{n-1}{2}}} \sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} \left(\left(\frac{n^2}{4N^2}\right) \sum_{i=1}^{n} j_i^2 + \frac{n^2}{2N^2} \sum_{i=1}^{n} j_i^3\right).$$

We also have

$$\sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} j_1^2 + ... + j_n^2 = 2n(b(N,n))^{n+1}$$

and

$$\sum_{j_{n-1}=-bh(N,n)}^{bh(N,n)} ... \sum_{j_1=-bh(N,n)}^{bh(N,n)} j_1^3 + ... + j_n^3 = 0.$$

It then follows that

$$\|f\|_\infty \frac{n^{\frac{n}{2}} b^{n+1}}{(2\pi N)^{\frac{n-1}{2}} n^{\frac{n+1}{2}}} \left(\frac{n^2}{4N^2}\right) nN^{\frac{n+1}{2}} = \frac{\|f\|_\infty C n^{\frac{5}{2}} b^{n+1}}{4(2\pi)^{\frac{n-1}{2}} N} < \epsilon$$

for large enough N. This proves the Theorem 1.8.1. That is, for each $\epsilon > 0$, there exists an $N_0 \in \mathbb{N}$ such that

$$|\frac{\sqrt{n}}{\left(\sqrt{2\pi}\right)^{n-1}} \int ... \int f\left(\frac{1}{\sqrt{n}}\sum_{n=1}^{n} o_i y_i\right) e^{\left(-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right)} dy_1...dy_{n-1}$$

$$-\sum_{\substack{k_1=1 \\ k_1+...+k_n=N}}^{N} ... \sum_{k_n=1}^{N} \frac{1}{n^N} \binom{N}{k_1, ..., k_n} f\left(\frac{\sum_{i=1}^{n} k_i o_i}{\sqrt{N}}\right) | < \epsilon$$

for all $N \geq N_0$. □

For future work, we would like to extend this result to a proof for $\mathbb{R}^n$ and, ultimately, for stochastic processes.

Chapter 2

The Estimation of the Concentration of Measure for Fractional Brownian Motion

2.1 Fractional Brownian Motion

The theory of stochastic processes is a foundational subject in the field of analysis. Brownian motion is the most fundamental and widely used stochastic process. A Brownian motion process $(W_t)_{t\in[0,T]}$ along with filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies

a. $W_0 = 0$.

b. $t \to W_t$ is continuous almost surely.

c. For all $0 \leq s < t \leq T,\ (W_t - W_s) \overset{d}{=} N(0, t-s)$.

d. For all $0 = t_0 < t_1 < ... < t_n \leq T,$ the increments

$$W_{t_1} - W_{t_0}, ..., W_{t_n} - W_{t_{n-1}}$$

are independent of one another.

For d, we may equivalently write that for $0 \leq r \leq s < t \leq T$, $W_t - W_s$ is indepdendent of the sigma algebra $\mathcal{F}_r$.

In 1900, Henri Poincare and his doctoral student Louis Bachlier modeled stock prices by a random walk, an approximation of Brownian motion, with the hope that this model would capture the randomness of stock prices. Einstein developed Brownian motion through his work on Avogadros number, the number of molecules in a mole of gas. Myron Scholes and Robert C. Merton developed the Black-Scholes option pricing theory in 1973 and won the Nobel Prize for their work in 1997.

A key feature of the Brownian motion process is independent time increments. Accordingly, this is reflected in the Black-Scholes model for returns on stock. However, this assumption is not always consistent with actual data. Instead, the past influences the present, creating dependent time increments.

Therefore, recent efforts to remedy this defect have led to consideration of fractional Brownian motion, in which the time increments depend on one another, and which Kolmogorov developed in his study of turbulence.

The strength of this dependence is controlled by the so-called Hurst index. The Hurst index H is a numerical parameter taking values in $[0, 1]$ where $H = \frac{1}{2}$ corresponds to the Brownian motion process. A fractional Brownian motion process $(W_t^H)_{t\in[0,T]}$ with Hurst index H has covariance structure

$$E(W_H(t)W_H(s)) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}).$$

The fractional Brownian motion process was developed by Kolmogorov in his study of turbulence. It is used today in fluid mechanics to model turbulent systems. To this day, fractional Brownian motion is used in fluid dynamics to model turbulent systems. In fact, Brownian motion and fractional Brownian motion have applications in modeling population growth, neuronal activity, genetic information, turbulent diffusion, radio-astronomy signals from stars, and the dynamics of satellites to mention just a few.

2.2 The Confidence Intervals for fBm

Theorem 2.2.1. *Let* $(W_t^H)_{t\in[0,\infty)}$ *be a fractional Brownian motion process with Hurst index* H*. Then,*

$$P(A_{N,\epsilon})$$

$$\geq \frac{|S_N|}{(2\pi)^{\frac{N}{2}}\sqrt{\prod_{i=1}^N \lambda_i}} \int_{\sqrt{(N-1)(1-\epsilon)}}^{\sqrt{(N+1)(1+\epsilon)}} e^{-\frac{1}{2}\rho^2 B}\left(\frac{\rho^{N-1}\sqrt{B}d\rho}{B^{\frac{N}{2}}}\right)$$

where

$$A_{N,\epsilon} = \{(N-1)(1-\epsilon) \leq \left(\sum_{i=1}^N (W_i - W_{i-1})^2\right) \leq (N-1)(1+\epsilon)\},$$

$$B = \frac{\rho^2}{N\pi}\sum_{i=1}^N \frac{1}{\lambda_i},$$

and (λ_i) *are the eigenvalues for the covariance matrix,* $M^{N,H}$*, for fractional Brownian noise.*

Proof. Consider the fractional Brownian motion process $(W_t^H)_{t\in[0,\infty)}$ with Hurst index H. By the Ergodic Theorem, we have

$$\frac{\sum_{i=1}^N (W_{t_i} - W_{t_{i-1}})^2}{N-1} \to \delta^{2H}$$

almost surely where $0 = t_0 < t_1 < ... < t_N$ and $|t_i - t_{i-1}| = \delta$ for each $1 \leq i \leq N$. By taking logs, we can solve for H. The Ergodic Theorem is discussed more in [2]. Let $t_i = i$ for each $1 \leq i \leq N$. We consider

$$P(1-\epsilon \leq \left(\frac{\sum_{i=1}^N (W_{t_i} - W_{t_{i-1}})^2}{(N-1)\delta^{2H}}\right) \leq 1+\epsilon)$$

$$\begin{aligned}
&= P((N-1)(1-\epsilon) \leq \left(\sum_{i=1}^N (W_i - W_{i-1})^2\right) \leq (N-1)(1+\epsilon)) \\
&= \int_{(N-1)(1-\epsilon)\leq \sum_{i=1}^N x_i^2 \leq (N-1)(1+\epsilon)} \cdots \int \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{detM}} e^{-\frac{1}{2}\bar{x}^T M^{-1}\bar{x}} dx_1 ... dx_N
\end{aligned}$$

Rotating to a basis of eigenvectors,

$$= \int \cdots \int_{(N-1)(1+\epsilon)\leq \sum_{i=1}^N x_i^2 \leq (N-1)(1+\epsilon)} \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\prod_{i=1}^N \lambda_i}} e^{-\frac{1}{2}(\sum_{i=1}^N \frac{x_i^2}{\lambda_i})} dx_1 ... dx_N$$

Converting to spherical coordinates,

$$= \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\prod_{i=1}^N \lambda_i}} \int \cdots \int e^{-\frac{1}{2}\sum_{i=1}^N \frac{x_i^2}{\lambda_i}} \rho^{N-1} \sin^{N-2}(\phi_1)\sin(\phi_2)^{N-3} ... \sin(\phi_{N-2}) d\rho d\phi_1 ... d\phi_N$$

where

$$\begin{aligned}
x_1 &= \rho\cos(\phi_1) \\
x_2 &= \rho\sin(\phi_1)\cos(\phi_2) \\
x_3 &= \rho\sin(\phi_1)\sin(\phi_2)\cos(\phi_3) \\
&\vdots \\
x_{N-1} &= \rho\sin(\phi_1)...\sin(\phi_{N-2})\cos(\phi_{N-1}) \\
x_N &= \rho\sin(\phi_1)...\sin(\phi_{N-1})
\end{aligned}$$

and $\sqrt{(N-1)(1+\epsilon)} \leq \rho \leq \sqrt{(N+1)(1+\epsilon)}$, $\phi_i \in [0, \pi)$ for $1 \leq i \leq N-2$, and $\phi_{N-1} \in [0, 2\pi)$. Multiplying and dividing by $|S_N| = \frac{2\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2})}$, we have

$$\frac{|S_N|}{|S_N|} \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\prod_{i=1}^N \lambda_i}} \int \cdots \int e^{-\frac{1}{2}\sum_{i=1}^N \frac{x_i^2}{\lambda_i}} \rho^{N-1} \sin^{N-2}(\phi_1)\sin(\phi_2)^{N-3} ... \sin(\phi_{N-2}) d\rho d\phi_1 ... d\phi_N$$

Since $f(x) = e^x$ is a convex function, by Jensen's Inequality we have

$$\frac{|S_N|}{|S_N|} \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\prod_{i=1}^N \lambda_i}} \int \cdots \int e^{-\frac{1}{2}\sum_{i=1}^N \frac{x_i^2}{\lambda_i}} \rho^{N-1} \sin^{N-2}(\phi_1)\sin(\phi_2)^{N-3} ... \sin(\phi_{N-2}) d\rho d\phi_1 ... d\phi_N$$

$$\geq \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\prod_{i=1}^N \lambda_i}} \int_{\sqrt{(N-1)(\epsilon-1)}}^{\sqrt{(N-1)(\epsilon+1)}} \rho^{N-1} e^{-\frac{1}{2}B} |S_N| d\rho$$

where

$$\begin{aligned}
B &= \frac{1}{|S_N|}\int_0^{2\pi}\int_0^{\pi}...\int_0^{\pi}\sum_{i=1}^{N}\frac{x_i^2}{\lambda_i}\sin^{N-2}(\phi_1)...\sin(\phi_{N-2})d\phi_1 d\phi_2...d\phi_N \\
&= \frac{1}{|S_N|}\sum_{i=1}^{N}\int_0^{2\pi}\int_0^{\pi}...\int_0^{\pi}\frac{x_i^2}{\lambda_i}\sin^{N-2}(\phi_1)...\sin(\phi_{N-2})d\phi_1 d\phi_2...d\phi_N \\
&= \frac{1}{|S_N|}\sum_{k=1}^{N} I_i.
\end{aligned}$$

Now, we will compute I_i and show that

$$I_i = \frac{2\pi\rho^2}{\lambda_i}\frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}$$

for all $1 \le i \le N$. Using $\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, we have

$$\begin{aligned}
I_1 &= \frac{1}{\lambda_1}\int_0^{2\pi}\int_0^{\pi}...\int_0^{\pi} x_1^2\sin^{N-2}(\phi_1)\sin^{N-3}(\phi_2)...\sin(\phi_{N-2})d\phi_1...d\phi_{N-1} \\
&= \frac{2\pi\rho^2}{\lambda_1}\int_0^{\pi}\cos^2(\phi_1)\sin^{N-2}(\phi_1)d\phi_1\int_0^{\pi}\sin^{N-3}(\phi_2)d\phi_2...\int_0^{\pi}\sin(\phi_{N-2})d\phi_{N-2} \\
&= \frac{2\pi\rho^2}{\lambda_1}\beta(\frac{N-1}{2},\frac{3}{2})\beta(\frac{N-2}{2},\frac{1}{2})...\beta(1,\frac{1}{2}) \\
&= \frac{2\pi\rho^2}{\lambda_1}\left(\frac{\Gamma(\frac{N-1}{2})\Gamma(\frac{3}{2})}{\Gamma(\frac{N+2}{2})}\right)\left(\frac{\Gamma(\frac{N-2}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N-1}{2})}\right)...\left(\frac{\Gamma(\frac{3}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{4}{2})}\right)\left(\frac{\Gamma(\frac{2}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})}\right) \\
&= \frac{2\pi\rho^2}{\lambda_1}\frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}.
\end{aligned}$$

$$\begin{aligned}
I_2 &= \frac{1}{\lambda_2}\int_0^{2\pi}\int_0^{\pi}...\int_0^{\pi} x_2^2\sin^{N-2}(\phi_1)\sin^{N-3}(\phi_2)...\sin(\phi_{N-2})d\phi_1...d\phi_{N-1} \\
&= \frac{2\pi\rho^2}{\lambda_2}\int_0^{\pi}\sin^2(\phi_1)\sin^{N-2}(\phi_1)d\phi_1\int_0^{\pi}\cos^2(\phi_2)\sin^{N-3}(\phi_2)d\phi_2\int_0^{\pi}\sin^{N-4}(\phi_3)d\phi_3... \\
&\qquad ...\int_0^{\pi}\sin(\phi_{N-2})d\phi_{N-2} \\
&= \frac{2\pi\rho^2}{\lambda_2}\int_0^{\pi}\sin^{N}(\phi_1)d\phi_1\int_0^{\pi}\cos^2(\phi_2)\sin^{N-3}(\phi_2)d\phi_2\int_0^{\pi}\sin^{N-4}(\phi_3)d\phi_3...
\end{aligned}$$

$$...\int_0^{\pi} \sin(\phi_{N-2}) d\phi_{N-2}$$

$$= \frac{2\pi\rho^2}{\lambda_2}\beta(\frac{N+1}{2},\frac{1}{2})\beta(\frac{N-2}{2},\frac{3}{2})\beta(\frac{N-3}{2},\frac{1}{2})...\beta(1,\frac{1}{2})$$

$$= \frac{2\pi\rho^2}{\lambda_2}\frac{\Gamma(\frac{N+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N+2}{2})}\frac{\Gamma(\frac{N-2}{2})\Gamma(\frac{3}{2})}{\Gamma(\frac{N+1}{2})}\frac{\Gamma(\frac{N-3}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N-2}{2})}...\frac{\Gamma(1)\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})}$$

$$= \frac{2\pi\rho^2}{\lambda_2}\frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}$$

$$= \frac{2\pi\rho^2}{\lambda_2}\frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}.$$

$$\vdots$$

$$\begin{aligned} I_i &= \frac{1}{\lambda_i}\int_0^{2\pi}\int_0^{\pi}...\int_0^{\pi} x_i^2 \sin^{N-2}(\phi_1)\sin^{N-3}(\phi_2)...\sin(\phi_{N-2}) d\phi_1...d\phi_{N-1} \\ &= \frac{2\pi\rho^2}{\lambda_i}\int_0^{\pi}\sin^2(\phi_1)\sin^{N-2}(\phi_1)d\phi_1\int_0^{\pi}\sin^2(\phi_2)\sin^{N-3}(\phi_2)d\phi_2... \end{aligned}$$

$$...\int_0^{\pi}\sin^2(\phi_{i-1})\sin^{N-i}(\phi_{i-1})d\phi_{i-1}\int_0^{\pi}\cos^2(\phi_i)\sin^{N-i-1}(\phi_i)d\phi_i\int_0^{\pi}\sin^{N-i-2}(\phi_{i+1})d\phi_{i+1}...$$

$$...\int_0^{\pi} \sin(\phi_{N-2}) d\phi_{N-2}$$

$$= \frac{2\pi\rho^2}{\lambda_i}\int_0^{\pi}\sin^N(\phi_1)d\phi_1\int_0^{\pi}\sin^{N-1}(\phi_2)d\phi_2...$$

$$...\int_0^{\pi}\sin^{N-i+2}(\phi_{i-1})d\phi_{i-1}\int_0^{\pi}\cos^2(\phi_i)\sin^{N-i-1}(\phi_i)d\phi_i\int_0^{\pi}\sin^{N-i-2}(\phi_{i+1})d\phi_{i+1}...$$

$$...\int_0^{\pi} \sin(\phi_{N-2}) d\phi_{N-2}$$

$$= \frac{2\pi\rho^2}{\lambda_i}\beta(\frac{N+1}{2},\frac{1}{2})\beta(\frac{N}{2},\frac{1}{2})...\beta(\frac{N-i+3}{2},\frac{1}{2})\beta(\frac{N-i}{2},\frac{3}{2})\beta(\frac{N-i-1}{2},\frac{1}{2})...\beta(1,\frac{1}{2})$$

$$= \frac{2\pi\rho^2}{\lambda_i}\frac{\Gamma(\frac{N+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N+2}{2})}\frac{\Gamma(\frac{N}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N+1}{2})}...\frac{\Gamma(\frac{N-i+3}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N-i+4}{2})}\frac{\Gamma(\frac{N-i}{2})\Gamma(\frac{3}{2})}{\Gamma(\frac{N-i+3}{2})}\frac{\Gamma(\frac{N-i-1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N-i}{2})}...\frac{\Gamma(1)\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})}$$

$$= \frac{2\pi\rho^2}{\lambda_i}\frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}.$$

$$\vdots$$

$$\begin{aligned}
I_{N-2} &= \frac{1}{\lambda_{N-2}}\int_0^{2\pi}\int_0^{\pi}...\int_0^{\pi} x_{N-1}^2 \sin^{N-2}(\phi_1)\sin^{N-3}(\phi_2)...\sin(\phi_{N-2})d\phi_1...d\phi_{N-1} \\
&= \frac{2\pi\rho^2}{\lambda_{N-2}}\int_0^{\pi}\sin^2(\phi_1)\sin^{N-2}(\phi_1)d\phi_1\int_0^{\pi}\sin^2(\phi_2)\sin^{N-3}(\phi_2)d\phi_3...
\end{aligned}$$

$$...\int_0^{\pi}\sin^2(\phi_{N-3})\sin^2(\phi_{N-3})d\phi_{N-3}\int_0^{\pi}\cos^2(\phi_{N-2})\sin(\phi_{N-2})d\phi_{N-2}$$

$$= \frac{2\pi\rho^2}{\lambda_{N-2}}\int_0^{\pi}\sin^{N}(\phi_1)d\phi_1\int_0^{\pi}\sin^{N-1}(\phi_2)d\phi_3...$$

$$...\int_0^{\pi}\sin^4(\phi_{N-3})d\phi_{N-3}\int_0^{\pi}\cos^2(\phi_{N-2})\sin(\phi_{N-2})d\phi_{N-2}$$

$$= \frac{2\pi\rho^2}{\lambda_{N-2}}\beta(\frac{N+1}{2},\frac{1}{2})\beta(\frac{N}{2},\frac{1}{2})...\beta(\frac{5}{2},\frac{1}{2})\beta(1,\frac{3}{2})$$

$$= \frac{2\pi\rho^2}{\lambda_{N-2}}\frac{\Gamma(\frac{N+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N+2}{2})}\frac{\Gamma(\frac{N}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N+1}{2})}...\frac{\Gamma(\frac{5}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{6}{2})}\frac{\Gamma(1)\Gamma(\frac{3}{2})}{\Gamma(\frac{5}{2})}$$

$$= \frac{2\pi\rho^2}{\lambda_{N-2}}\frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}.$$

$$\begin{aligned}
I_{N-1} &= \frac{1}{\lambda_{N-1}}\int_0^{2\pi}\int_0^{\pi}...\int_0^{\pi} x_{N-1}^2 \sin^{N-2}(\phi_1)\sin^{N-3}(\phi_2)...\sin(\phi_{N-2})d\phi_1...d\phi_{N-1} \\
&= \frac{\rho^2}{\lambda_{N-1}}\int_0^{\pi}\sin^2(\phi_1)\sin^{N-2}(\phi_1)d\phi_1\int_0^{\pi}\sin^2(\phi_2)\sin^{N-3}(\phi_2)d\phi_2...
\end{aligned}$$

$$...\int_0^{\pi}\sin^2(\phi_{N-2})\sin(\phi_{N-2})d\phi_{N-2}\int_0^{2\pi}\cos^2(\phi_{N-1})d\phi_{N-1}$$

$$= \frac{\pi\rho^2}{\lambda_{N-1}}\int_0^{\pi}\sin^{N}(\phi_1)d\phi_1\int_0^{\pi}\sin^{N-1}(\phi_2)d\phi_2...\int_0^{\pi}\sin^3(\phi_{N-2})d\phi_{N-2}$$

$$= \frac{\pi\rho^2}{\lambda_{N-1}}\beta(\frac{N+1}{2},\frac{1}{2})\beta(\frac{N}{2},\frac{1}{2})...\beta(\frac{4}{2},\frac{1}{2})$$

$$= \frac{2\pi\rho^2}{\lambda_{N-1}}\left(\frac{1}{2}\frac{\Gamma(\frac{N+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N+2}{2})}\frac{\Gamma(\frac{N}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{N+1}{2})}...\frac{\Gamma(\frac{4}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{5}{2})}\right)$$

$$= \frac{2\pi\rho^2}{\lambda_{N-1}} \left(\frac{1}{2} \frac{(\Gamma(\frac{1}{2}))^{N-2}\Gamma(2)}{\Gamma(\frac{N+2}{2})} \right) = \frac{2\pi\rho^2}{\lambda_{N-1}} \left(\frac{1}{2}\right) \left(\frac{\Gamma(\frac{3}{2})\Gamma(\frac{1}{2})^{N-3}}{\Gamma(\frac{N+2}{2})} \right) \left(\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})} \right)$$

since $\Gamma(z+1) = z\Gamma(z)$,

$$= \frac{2\pi\rho^2}{\lambda_{N-1}} \left(\frac{1}{2}\right) \left(\frac{\Gamma(\frac{3}{2})\Gamma(\frac{1}{2})^{N-3}}{\Gamma(\frac{N+2}{2})} \right) \left(\frac{\Gamma(\frac{1}{2})}{\frac{1}{2}\Gamma(\frac{1}{2})} \right)$$

$$= \frac{2\pi\rho^2}{\lambda_{N-1}} \frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}.$$

$$\begin{aligned}
I_N &= \frac{1}{\lambda_N} \int_0^{2\pi} \int_0^{\pi} ... \int_0^{\pi} x_N^2 \sin^{N-2}(\phi_1)\sin^{N-3}(\phi_2)...\sin(\phi_{N-2}) d\phi_1 ... d\phi_{N-1} \\
&= \frac{\rho^2}{\lambda_N} \int_0^{\pi} \sin^2(\phi_1)\sin^{N-2}(\phi_1) d\phi_1 \int_0^{\pi} \sin^2(\phi_2)\sin^{N-3}(\phi_2) d\phi_2 ... \\
&\quad ... \int_0^{\pi} \sin^2(\phi_{N-2})\sin(\phi_{N-2}) d\phi_{N-2} \int_0^{2\pi} \sin^2(\phi_{N-1}) d\phi_{N-1} \\
&= \frac{\pi\rho^2}{\lambda_N} \int_0^{\pi} \sin^N(\phi_1) d\phi_1 \int_0^{\pi} \sin^{N-1}(\phi_2) d\phi_2 ... \int_0^{\pi} \sin^3(\phi_{N-2}) d\phi_{N-2} \\
&= \frac{\pi\rho^2}{\lambda_N} \beta(\frac{N+1}{2}, \frac{1}{2}) \beta(\frac{N}{2}, \frac{1}{2}) ... \beta(2, \frac{1}{2}). \\
&= \frac{2\pi\rho^2}{\lambda_N} \frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}\Gamma(1)}{\Gamma(\frac{N+2}{2})}.
\end{aligned}$$

It then follows that

$$I_i = \frac{2\pi\rho^2}{\lambda_i} \frac{\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}}{\Gamma(\frac{N+2}{2})}$$

for all $1 \leq i \leq N$ and thus

$$\begin{aligned}
B &= \frac{2\pi\rho^2\Gamma(\frac{3}{2})(\Gamma(\frac{1}{2}))^{N-3}}{N\pi^{\frac{N}{2}}} \sum_{i=1}^{N} \frac{1}{\lambda_i} \\
&= \frac{\rho^2}{\pi N} \sum_{i=1}^{N} \frac{1}{\lambda_i}.
\end{aligned}$$

Therefore, we have

$$P((N-1)(1-\epsilon) \leq \left(\sum_{i=1}^{N}(W_i - W_{i-1})^2\right) \leq (N-1)(1+\epsilon))$$

$$\geq \frac{|S_N|}{(2\pi)^{\frac{N}{2}}\sqrt{\prod_{i=1}^{N}\lambda_i}} \int_{\sqrt{(N-1)(1-\epsilon)}}^{\sqrt{(N+1)(1+\epsilon)}} e^{-\frac{1}{2}\rho^2 B}\left(\frac{\rho^{N-1}\sqrt{B}d\rho}{B^{\frac{N}{2}}}\right)$$

where

$$B = \frac{\rho^2}{N\pi}\sum_{i=1}^{N}\frac{1}{\lambda_i}.$$

□

Thus, all of the information in the bounds is obtained by understanding the spectrum of the covariance operator.

2.3 The Spectrum for the Covariance Matrix of fBm Increments

Now, to determine the confidence intervals for fractional Brownian motion, we study the spectrum of the covariance operator.

Theorem 2.3.1. *The maximum eigenvalue for the covariance matrix for fractional Brownian noise of dimension N and Hurst index H has a rate of N^{2H-1} for $H \in (\frac{1}{2}, 1)$. For $H \in (0, \frac{1}{2})$, an upper bound for the largest eigenvalue is $\frac{3}{2}$.*

Proof. The covariance function for a fractional Brownian motion process $(W_H(t))_{t\in[0,T]}$ is given by

$$E(W_H(t)W_H(s)) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}).$$

Therefore, the matrix $M^{N,H}$ has entries $M_{i,j}^{N,H} = \frac{1}{2}(|i-j+1|^{2H} + |i-j-1|^{2H} - 2|i-j|^{2H})$ for some $H \in (0,1)$. By Brauer's Theorem, the spectrum of a matrix is contained in the union of circles of radius equal to the absolute value of the sum of the rows without the diagonal entry and center equal to the diagonal entry. For the sum of row k without the diagonal entry, we have

$$\sum_{m=0, m\neq k}^{N} \frac{1}{2}(|m-k+1|^{2H} + |m-k-1|^{2H} - 2|m-k|^{2H})$$

as a telescoping sum,

$$= \frac{1}{2}((k+1)^{2H} - k^{2H} + (N-k+1)^{2H} - (N-k)^{2H}) + (-1).$$

Now, we will find the maximum value of this sum over all k. Let

$$f(k) = \frac{1}{2}((k+1)^{2H} - k^{2H} + (N-k+1)^{2H} - (N-k)^{2H}) + (-1).$$

Then, we have

$$f'(k) = \frac{1}{2}(2H(k+1)^{2H-1} - 2Hk^{2H-1} - 2H(n-k+1)^{2H-1} + 2H(N-k)^{2H-1}).$$

Setting $f'(k) = 0$, we obtain

$$(k+1)^{2H-1} + (N-k)^{2H-1} = k^{2H-1} + (N-k+1)^{2H-1},$$

which has $k = \frac{N}{2}$ as a solution. For the second derivative test, we have

$$f''(k) = H(2H-1)((k+1)^{2H-2} - k^{2H-2} + (N-k+1)^{2H-2} - (N-k)^{2H-2}).$$

Since $(2H-2) < 0$ for each $H \in (0,1)$, we have

$$((k+1)^{2H-2} - k^{2H-2}) \leq 0$$

and

$$((N-k+1)^{2H-2} - (N-k)^{2H-2}) \leq 0$$

for $k > 0$. Therefore, for $H \in (0, \frac{1}{2})$, we have $(2H-1) < 0$ and thus $f"(k) > 0$, so $f(0)$ or $f(N)$ is a maximum. Thus, for $H \in (0, \frac{1}{2})$, we have

$$f(0) = f(N) = \frac{1}{2}((N+1)^{2H} - N^{2H} + 1) + (-1)$$

$$= \frac{N^{2H}}{2}(1+\frac{1}{N})^{2H} - \frac{N^{2H}}{2} - \frac{1}{2}$$
$$= \frac{N^{2H}}{2}(1+\frac{2H}{N} + O(\frac{1}{N^2})) - \frac{N^{2H}}{2} - \frac{1}{2}$$
$$= HN^{2H-1} - \frac{1}{2} + O(N^{2H-2}).$$

Thus, $\frac{1}{2} - HN^{2H-1} + 1$ is an upper bound for the largest eigenvalue. Similarly, for $H \in (\frac{1}{2}, 1)$, we have $(2H-1) > 0$ and thus $f''(k) < 0$, so $f(\frac{N}{2})$ is a maximum. Thus, for $H \in (\frac{1}{2}, 1)$

$$f(\frac{N}{2}) + 1 = ((\frac{N}{2}+1)^{2H} - (\frac{N}{2})^{2H})$$

gives us an upper bound for the largest eigenvalue. We have

$$((\frac{N}{2}+1)^{2H} - (\frac{N}{2})^{2H}) = (\frac{N}{2})^{2H}(1+(\frac{2}{N}))^{2H} - (\frac{N}{2})^{2H}.$$

Using Taylor series approximation, we have

$$((\frac{N}{2}+1)^{2H} - (\frac{N}{2})^{2H}) = (\frac{N}{2})^{2H}(1+2H(\frac{2}{N}) + O(\frac{1}{N^2})) - (\frac{N}{2})^{2H}$$
$$= \frac{4H}{4^H}N^{2H-1} + O(N^{2H-2}).$$

Since $H \in (\frac{1}{2}, 1)$, then $O(N^{2H-2}) \to 0$ as $N \to \infty$. Now, we will find a lower bound for the largest eigenvalue of M. To do this, we will consider $< Mx, x >$ where

$$x = \frac{1}{\sqrt{N}}(11...1)^T.$$

Then, we have

$$< Mx, x >= \frac{1}{N}\sum_{i=0}^{N}\sum_{j=0}^{N}\frac{1}{2}(|i-j+1|^{2H} + |i-j-1|^{2H} - 2|i-j|^{2H})$$

As a telescoping sum,

$$= \frac{1}{N}(N+1)^{2H} = \frac{1}{N}((N^{2H})(1+\frac{1}{N})^{2H})$$

Using Taylor series approximation,

$$= N^{2H-1}(1+o(\frac{1}{N})) = N^{2H-1} + O(N^{2H-2}).$$

Since $H \in (\frac{1}{2}, 1)$, then $O(N^{2H-2}) \to 0$ as $N \to \infty$. Now, we will show that the upper and lower bounds for the matrix M are sufficiently close to one another. The largest eigenvalue $\lambda^{N,H}$ corresponding to matrix $M^{N,H}$ must satisfy

$$N^{2H-1} \leq \lambda^{N,H} \leq \frac{4H}{4^H} N^{2H-1}.$$

To find the maximum value of $\frac{4H}{4^H}$, let

$$g(H) = \frac{4H}{4^H}.$$

Then, we have

$$g'(H) = \frac{4 \cdot 2^{2H} - 4H \cdot 2^{2H} \cdot 2ln(2)}{2^{4H}}.$$

Setting $g'(H) = 0$, we obtain

$$H = \frac{1}{2ln(2)}.$$

Then, we have maximum value

$$1.06 < g(\frac{1}{2ln(2)}) = \frac{2}{ln(2) \cdot 2^{\frac{1}{ln(2)}}} < 1.062.$$

Therefore, we have largest eigenvalue $\lambda_{N,H}$ of matrix $M^{N,H}$ with rate equal to $a(H)N^{2H-1}$ where $1 \leq a(H) < 1.062$. □

Thus, we have the rate for the largest eigenvalue for M. For future work, we would like to continue our analysis of the spectrum of M to find tight bounds for the concentration of measure for fBm.

Chapter 3

Appendix

3.1 Appendix A.

A sequence of random variables (X_n) on a probability space $(\Omega, \mathcal{F}, P)$ is said to converge "almost surely" to a random variable X provided that

$$P\left(\{\omega \in \Omega : X_n(\omega) \nrightarrow X(\omega)\}\right) = 0.$$

This sequence of random variables converges to X "in probability" provided that for each $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}\right) = 0.$$

A sequence of random variables (X_n) is said to converge to a random variable X in the L^pnorm provided that

$$\lim_{n\to\infty} \|X_n - X\|_p = 0$$

where $\|X\|_p = \left(\int_\Omega X(\omega)^p dP(\omega)\right)^{\frac{1}{p}}$.

3.2 Appendix B.

The following notes on Laplace are from [4]. Laplace first became interested in the probabilities of randomly distributed sums of angles of inclination. In 1776, he published a paper determining the probability that the sums of angles of inclination of comet orbits were within given limits. In 1781, Laplace developed a general method for calculating these probabilities based on convolutions of density functions. In the most simple case, Laplace considered n variables with the same rectangular distribution between 0 and h. Then, the probability that the sum of those variables is between a and b is given by

$$P = \frac{1}{h^n n!}\left(\sum_{i=0}^{N}\binom{n}{i}(-1)^i(b-ih)^n - \sum_{i=0}^{M}\binom{n}{i}(-1)^i(a-ih)^n\right)$$

where $N = \min(n, \lfloor \frac{b}{n} \rfloor)$ and $M = \min(n, \lfloor \frac{a}{n} \rfloor)$. However, this formula was too intricate for direct numerical computations.

Laplace was particularly interested in approximating integrals by studying functions with a sharp peak. For these functions, the main contribution to the integral is contained in some small interval around the value at which the maximum occurrs. He approximated these functions by $f(a)e^{-\alpha(x-a)^2}$ for a function f whose maximum occurs at $x = a$. One can easily see the connection between this approximation and the density for the normal distribution. An example of an integral that Laplace studied is the gamma function

$$\Gamma(s+1) = \int_0^\infty e^{-x}x^s dx.$$

Letting $x = z + s$,

$$\Gamma(s+1) = \int_{-s}^\infty e^{-(z+s)}(z+s)^s dz.$$

To approximate this integral, Laplace recognized that the integrand has maximum value at $M = e^{-s}s^s$ attained at $x = s$, or equivalently, at $z = 0$. Then, Laplace set

$$e^{-(z+s)}(z+s)^s = e^{-s}e^{-z}(z+s)^s = Me^{-t^2 z}$$

and expanded $t^2 = -\frac{1}{z}\log(e^{-z}(1+\frac{z}{s})^s)$ in a power series about z. He also expanded z in a power series about t. After transforming the variable of integration from z to

t, he obtained

$$\begin{aligned}\Gamma(s+1) &= M\int_{-\infty}^{\infty} e^{-t^2}\sqrt{2s}\left(1+\frac{4t}{3\sqrt{2s}}+\frac{t^2}{6s}+...\right)dt\\ &= s^{s+\frac{1}{2}}e^{-s}\sqrt{2\pi}\left(1+\frac{1}{12s}+\frac{1}{288s^2}+...\right).\end{aligned}$$

This computation provides an example of Laplace's method for approximation. For many problems, this method worked well in establishing a close approximation. However, this method did not work well for the case of sums of random variables. Therefore, another method was needed.

In his next approach, Laplace was able to compute probabilities for sums of independent random variables. As an example, consider random variables $X_1, ..., X_n$ with mean 0 which take values $\frac{k}{m}$, $m \in \mathbb{N}$, $k = -m, -m+1, ..., m-1, m$ with probabilities p_k. Then, let P_j denote the probability that $\sum_{l=1}^{n} X_l$ has the value $\frac{j}{m}, -nm \leq j \leq nm$. Laplace used the generating function $T(t) = \sum_{k=-m}^{m} p_k t^k$. By independence of the random variables, P_j is equal to the coefficient of t^j in the product $(T(t))^n$. As the execution of this method is extremely complicated, dating back to de Moivre, Laplace employed the trick of letting $t = e^{ix}$. As a result of

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itx}e^{isx}dx = \delta_{t,s}$$

it follows that

$$P(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ijx}\left(\sum_{k=-m}^{m} p_k e^{ikx}\right)^n dx.$$

Expanding e^{ikx} in power series,

$$P(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ijx}\left(\sum_{k=-m}^{m} p_k\left(1+ikx-\frac{k^2x^2}{2}-...\right)\right)^n dx.$$

As the random variables (X_l) have mean 0, then we have $\sum_{k=-m}^{m} p_k k = 0$ and $\sum_{k=-m}^{m} p_k k^2 = m^2\sigma^2$. Making these substitutions, we have

$$P(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ijx}\left(1-\frac{m^2\sigma^2x^2}{2}-iAx^3+...\right)dx$$

where A is a constant depending on $\sum_{k=-m}^{m} p_k k^3$. Expanding

$$\log\left(1-\frac{m^2\sigma^2x^2}{2}-iAx^3+...\right)^n := \log(z)$$

in power series, we have

$$\log(z) = -\frac{m^2\sigma^2nx^2}{2} - iAnx^3 + ...$$

which gives us

$$\begin{aligned} z &= e^{-\frac{m^2\sigma^2nx^2}{2}-iAnx^3+...} \\ &= e^{-\frac{m^2\sigma^2nx^2}{2}}\left(1-iAnx^3+...\right). \end{aligned}$$

Letting $y = x\sqrt{n}$, we have

$$P(j) = \frac{1}{2\pi\sqrt{n}}\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} e^{-ij\frac{y}{\sqrt{n}}}e^{-\frac{m^2\sigma^2y^2}{2}}\left(1-iAy^3+...\right)dy.$$

For very large n, we have

$$P(j) = \frac{1}{2\pi\sqrt{n}}\int_{-\infty}^{\infty} e^{-ij\frac{y}{\sqrt{n}}}e^{-\frac{m^2\sigma^2y^2}{2}}dy.$$

Laplace showed in different ways that this integral is equal to

$$\frac{1}{m\sigma\sqrt{2\pi n}}e^{\frac{-j^2}{2m^2\sigma^2n}}(*).$$

Summing up $(*)$ for $\frac{j}{m} \in [r_1\sqrt{n}, r_2\sqrt{n}]$, we have

$$P\left(r_1\sqrt{n} \le \sum X_l \le r_2\sqrt{n}\right) \approx \sum_{j\in[mr_1\sqrt{n},mr_2\sqrt{n}]} \frac{1}{m\sigma\sqrt{2\pi n}}e^{\frac{-j^2}{2m^2\sigma^2n}},$$

approximating by an integral with $\Delta x = \frac{1}{\sqrt{n}}$,

$$\approx \int_{mr_1}^{mr_2}\frac{1}{m\sigma\sqrt{2\pi}}e^{\frac{-j^2}{2m^2\sigma^2}}dx$$

$$= \int_{r_1}^{r_2}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}dx.$$

This gives us a special case of the CLT by establishing the probability that a sum of random variables is within given limits. As he starts by considering discrete random variables, he later extends to the case with m "infinitely large." Laplace never proved the general CLT that we use today. Instead, he considered the approximate probabilities involving linear combinations of observed errors. His most general version of the CLT is the following: Let $\epsilon_1, ..., \epsilon_n$ be independent observation errors with mean μ and variance σ^2. Let $\lambda_1, ..., \lambda_n$ be constant multipliers and $a > 0$. Then, we have

$$P\left(|\sum_{j=1}^{n} \lambda_j(\epsilon_j - \mu)| \leq a\sqrt{\sum_{j=1}^{n} \lambda_j^2}\right) \approx \frac{2}{\sigma\sqrt{2\pi}} \int_0^a e^{-\frac{x^2}{2\sigma^2}} dx.$$

3.3 Appendix C.

The following notes on Poisson are from [4]. Unlike Laplace, Poisson started with sums of random variables from the beginning. By considering random variables (X_n) with density functions f_n and by letting $S_s = X_1 + ... + X_s$, Poisson established the formula

$$P\left(c - \epsilon \leq S_s \leq c + \epsilon\right) = \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\prod_{n=1}^{s} \int_a^b f_n(x) e^{i\alpha x} dx\right) e^{i\alpha c} \sin(\epsilon\alpha) \frac{d\alpha}{\alpha}$$

However, the justification for this formula was incomplete. He then considered the special case with $s = 1$. Changing the order of integration, he obtained the formula

$$P\left(c - \epsilon \leq X_1 \leq c + \epsilon\right) = \frac{1}{\pi} \int_a^b \int_{-\infty}^{\infty} e^{i\alpha(x-c)} \sin(\epsilon\alpha) \frac{d\alpha}{\alpha} f_1(x) dx.$$

By using the formula

$$\int_0^{\infty} \frac{\sin(kx)}{x} dx = \frac{\pi}{2}, k > 0$$

he obtained

$$\int_{-\infty}^{\infty} e^{i\alpha(x-c)} \sin(\epsilon\alpha) \frac{d\alpha}{\alpha} = \begin{cases} \pi & x \in [c - \epsilon, c + \epsilon] \\ 0 & x \notin [[c - \epsilon, c + \epsilon]. \end{cases}$$

In order to establish this formula, Poisson required

$$P(c-\epsilon \leq X_1 \leq c+\epsilon) = \int_{c-\epsilon}^{c+\epsilon} f_1(x)dx.$$

To extend to the general case, Poisson set

$$\int_a^b f_n(x)\cos(\alpha x)dx := \rho_n \cos(\phi_n)$$

$$\int_a^b f_n(x)\sin(\alpha x)dx := \rho_n \sin(\phi_n)$$

where $R := \rho_1 ... \rho_s$ and $\psi := \phi_1 + ... + \phi_s$. Since $R(-\alpha) = R(\alpha)$ and $\psi(-\alpha) = -\psi(\alpha)$, he showed that

$$P(c-\epsilon \leq S_s \leq c+\epsilon) = \frac{2}{\pi}\int_0^\infty R\cos(\psi - c\alpha)\sin(\epsilon\alpha)\frac{d\alpha}{\alpha}.$$

This formula gives us the probability that a sum of the large number of random variables is within a given limit. While Poisson's work on the CLT was based on the work of Laplace, Poisson's strict mathematical analysis led to more rigorous treatment of the CLT. Poisson's version of the CLT can be summarized as follows: Let $X_1, ..., X_s$ be random variables with densities that decrease sufficiently fast as their arguments tend to $\pm\infty$. Suppose that for the absolute values $\rho_n(\alpha)$ of the characteristic function X_n there exists a function $r(\alpha)$ independent of n with $0 \leq r(\alpha) < 1$ for each $\alpha \neq 0$ such that

$$\rho_n(\alpha) \leq r(\alpha).$$

Then, for arbitrary γ,γ',

$$P\left(\gamma \leq \frac{\sum_{n=1}^s (X_n - EX_n)}{\sqrt{2\sum_{n=1}^s VarX_n}} \leq \gamma'\right) \approx \frac{1}{\sqrt{\pi}}\int_\gamma^{\gamma'} e^{-u^2}du.$$

To investigate the validity of his work, Poisson considered a counterexample for random variables with density function equal to

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

Then, we have the probability that the sum of the random variables is within a fixed limit given by

$$P(c-\epsilon \leq \sum X_n \leq c+\epsilon) = \frac{1}{\pi}\text{arctan}\left(\frac{\epsilon s}{s^2+c^2-\epsilon^2}\right)$$

Therefore, the probability is not approximated by the normal distribution for large s. Next, Poisson considered linear combinations of identically distributed errors $\sum \gamma_n \epsilon_n$ which satisfy

$$f(x) = e^{-2|x|}.$$

These linear combinations of errors satisfy

$$P(-c \leq \sum \gamma_n \epsilon_n \leq c) = \frac{1-e^{-2c}}{1+e^{2c}}$$

if $\gamma_n = \frac{1}{n}$, and

$$P(-c \leq \sum \gamma_n \epsilon_n \leq c) = 1 - \frac{4}{\pi}\text{arctan}(e^{-2c})$$

if $\gamma_n = \frac{1}{2n-1}$. In the first example, we have

$$\rho_1(\alpha)...\rho_s(\alpha) = \frac{1}{(1+\frac{\alpha^2}{4})(1+\frac{\alpha^2}{4\cdot 4})...(1+\frac{\alpha^2}{4\cdot s^2})} \to \frac{\pi\alpha}{e^{\frac{1}{2}\pi\alpha} - e^{-\frac{1}{2}\pi\alpha}}.$$

In the second example, we have

$$\rho_1(\alpha)...\rho_s(\alpha) = \frac{1}{(1+\frac{\alpha^2}{4})(1+\frac{\alpha^2}{4\cdot 9})...(1+\frac{\alpha^2}{4\cdot(2s-1)^2})} \to \frac{2}{e^{\frac{\pi\alpha}{4}} - e^{-\frac{\pi\alpha}{4}}}.$$

Poisson's earlier version of the CLT led the way for a more rigorous treatment. Letting

$$\rho := \rho_1 = \sqrt{(\int_a^b f_1(x)\text{cos}(\alpha x)dx)^2 + (\int_a^b f_1(x)\text{sin}(\alpha x)dx)^2}$$

and $\phi := \phi_1$, Poisson showed that

$$P(c-\epsilon \leq S_s \leq c+\epsilon) = \frac{2}{\pi}\int_0^\infty \rho^s \text{cos}(s\phi - c\alpha)\text{sin}(\epsilon\alpha)\frac{d\alpha}{\alpha}.$$

For "infinitely small" α, he deduced that

$$\rho^s = \begin{cases} (1-h^2\alpha^2)^s & \alpha \text{ infinitely small} \\ 0 & \text{otherwise} \end{cases}$$

where

$$h^2 := \frac{1}{2}\left(\int_a^b x^2 f_1(x)dx - \left(\int_a^b x f_1(x)dx\right)^2\right).$$

For "infinitely large" s and large but finite Y,

$$P(c-\epsilon \le S_s \le c+\epsilon) \approx \frac{2}{\pi}\int_0^Y e^{-h^2y^2}\cos[(ks-c)\frac{y}{\sqrt{s}}]\sin(\frac{\epsilon y}{\sqrt{s}})\frac{dy}{y}+$$

$$+\frac{2}{\pi}\int_{\frac{Y}{\sqrt{s}}}^{\infty} \rho^s\cos(s\phi - c\alpha)\sin(\epsilon\alpha)\frac{d\alpha}{\alpha}.$$

Poisson observed that

$$\frac{2}{\pi}\int_{\frac{Y}{\sqrt{s}}}^{\infty} \rho^s cos(s\phi - c\alpha) sin(\epsilon\alpha)\frac{d\alpha}{\alpha} \approx 0.$$

Using the equality,

$$\frac{1}{y}\cos[(ks-c)\frac{y}{\sqrt{s}}]\sin(\frac{\epsilon y}{\sqrt{s}}) = \frac{1}{\pi\sqrt{s}}\int_{-\epsilon}^{\epsilon}\cos[(ks-c+z)\frac{y}{\sqrt{s}}]dz$$

we have

$$\frac{2}{\pi}\int_0^Y e^{-h^2y^2}\cos[(ks-c)\frac{y}{\sqrt{s}}]\sin(\frac{\epsilon y}{\sqrt{s}})\frac{dy}{y}$$

$$= \frac{1}{\pi\sqrt{s}}\int_{-\epsilon}^{\epsilon}\left(\int_0^{\infty} e^{-h^2y^2}\cos[(ks-c+z)\frac{y}{\sqrt{s}}]dy\right)dz$$

Thus, we have

$$P(c-\epsilon \le S_s \le c+\epsilon) \approx \frac{1}{2h\sqrt{\pi s}}\int_{-\epsilon}^{\epsilon} e^{-\frac{(ks-c+z)^2}{4h^2 s}}dz.$$

Setting $c = ks$ and $\epsilon = 2hr\sqrt{s}$, he obtained the following version of the CLT

$$P(ks - 2hr\sqrt{s} \le S_s \le ks + 2hr\sqrt{s}) \approx \frac{2}{\sqrt{\pi}}\int_0^r e^{-t^2}dt.$$

3.4 Appendix D.

For an example, we will show that $\frac{3}{8}$ has binary expansion equal to $(0, 1, 1, 0, 0, 0, ...)$ as follows: We first divide $[0, 1)$ into two intervals of equal length, $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$. Since $\frac{3}{8}$ is in the left interval $[0, \frac{1}{2})$, we write 0 as the first digit in the binary expansion. Next, we divide $[0, \frac{1}{2})$ into two intervals of equal length, $[0, \frac{1}{4})$ and $[\frac{1}{4}, \frac{1}{2})$. Since $\frac{3}{8}$ is in the right interval $[\frac{1}{4}, \frac{1}{2})$, we write 1 as the second digit. Continuing the process, we divide $[\frac{1}{4}, \frac{1}{2})$ into two intervals of equal length, $[\frac{1}{4}, \frac{3}{8})$ and $[\frac{3}{8}, \frac{1}{2})$. Since $\frac{3}{8}$ is in the right interval $[\frac{3}{8}, \frac{1}{2})$, then we write 1 as the third digit in the binary expansion for $\frac{3}{8}$. One can see that since $\frac{3}{8}$ is the left endpoint of the interval $[\frac{3}{8}, \frac{1}{2})$, $\frac{3}{8}$ will always be in the left interval following each of the proceeding subdivisions. Therefore, $\frac{3}{8}$ has binary expansion equal to $(0, 1, 1, 0, 0, 0, ...)$. We can check that $\frac{0}{2} + \frac{1}{4} + \frac{1}{8} = \frac{3}{8}$.

Bibliography

[1] Andrew R. Barron, *Entropy and the central limit theorem*, The Annals of Probability **14** (1986), no. 1, 336–342.

[2] Patrick Billingsley, *Probability and Measure*, John Wiley and Sons, Inc., 1995.

[3] Rick Durrett, *Probability: Theory and Examples*, Cambridge University Press, 2010.

[4] Hans Fischer, *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*, Springer, 2011.

[5] G. G. Lorentz, *Bernstein Polynomials*, American Mathematical Society, New York, 1986.

[6] Walter Rudin, *Functional Analysis*, McGraw-Hill Science, 1991.

[7] Bengt von Bahr, *On the convergence of moments in the central limit theorem*, The Annals of Mathematical Statistics **36** (1965), no. 3, 808–818.

[8] David Walnut, *An Introduction to Wavelet Analysis*, Birkhauser, 2004.

Curriculum Vitae

Patricia Garmirian

Email: pmg208@lehigh.edu
Cell phone: (484) 335-3595

Department of Mathematics
Lehigh University
14 E. Packer Ave.
Bethlehem, PA 18015

Citizenship: United States.

Research interests:

Probability and Stochastic Processes.

Education:

Lehigh University — Bethlehem, PA
September 2008-July 2013
Ph.D., Mathematics, July 2013.
Thesis *The Central Limit Theorem and the Estimation for the Concentration of Measure for Fractional Brownian Motion* supervised by Vladimir Dobric.

Indiana University — Bloomington, IN
September 2005-May 2006
Graduate Student

Denison University — Granville, OH
September 2001-May 2005
B.S., Mathematics, May 2005

Lehigh University Bethlehem, PA

August 2008-May 2013

Teaching Assistant. Teaching assistant for various levels of calculus, including multivariate, ran recitation sessions once a week for four sections per semester, graded papers/homework and exams, held office hours.

Grants and awards:

Lehigh University Bethlehem, PA

June 2012-August 2012

Strohl Summer Research Fellowship.

Presentations:

GSIMS, Lehigh University, 2010, 2011, 2012, 2013.

Joint AMS/AMA Meeting, 2005.