CENTRAL LIMIT THEOREM VIA MOMENTS

DUNCAN HAYSTEAD

1. Introduction

One of the most important result in Modern Probability theory is the Central Limit Theorem, which states that fluctuations of the average of an iid sequence of random variables are asymptotically gaussian (see Theorem 4.1 for the exact technical statement). The result was first proven Bernoulli random variables using asymptotics of $N!$ by De Moivre (1733) and Laplace (1812). However, it was only in the early 20th century, when Lyapunov (1901) was able to extend to all random variables having finite second moment. Polya called this the 'central limit theorem' due to its importance in probability theory. Since then there have many extensions of this result, and very large area of probability theory deals with this Gaussian distribution. This 'bell curve' (see figure 1) is quite ubiquitous and in fact is the basis for much of classical statistics.

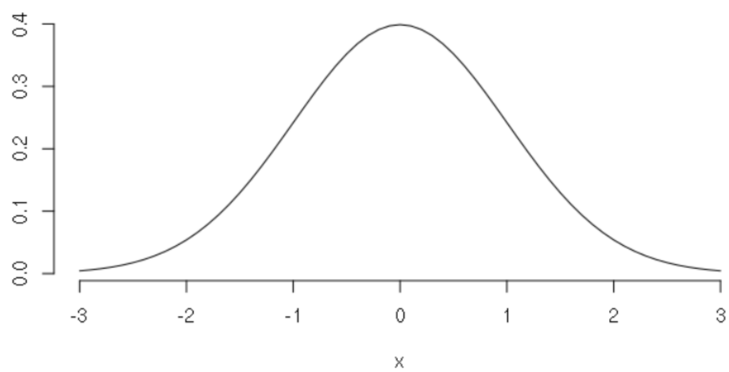

Figure 1. The Gaussian bell curve

In this note, we provide a nonstandard proof of the central limit theorem that relies on moment calculations and is related to the classical moment problem.

The moment problem in probability theory studies the map $\mathcal{S}$ taking a random variable X to its moment sequence $\{s_k\}_{k\geq 0}$,

$$s_k(\mu(X)) = \int x^k d\mu(x) = \mathbb{E}(X^k) \tag{1.1}$$

wher μ is the probability measure associated with X. Two fundamental questions in the moment problem are existence and determinacy, i.e., for which random variable X does the moment sequence exist and for which random variable X is the moment sequence uniquely characterized by X?

The classical moment problem originated in the 1880s, and reached a definitive state by the end of the 1930s. One of the original sources of motivation came from probability theory, where it is important to have verifiable sufficient conditions for determinacy. Determinacy is also closely related to several problems in classical analysis, particularly, to the study of the map taking a (germ of a) smooth function f to the sequence of its Taylor coefficients $(f^{(k)}(0))_{k\geq 0}$. Existence in the moment problem is a prototype of the problem of extension of a positive functional, and it gave the impetus for the development of several functional-analytic tools.

In this paper, we study some of the results on determinacy the classical moment problem including the Denjoy-Carleman theorem and its corollary Carleman's condition. The results are stated in Section 3. Using these results, we give a combinatorial proof of the Central Limit Theorem in Section 4. The techniques of this combinatorial proof is of independent interest as it is also used in proving Wigner's semicircle law, one of the most of important result in Random Matrix Theory.

2. Preliminaries

In this section, we recall some basic facts about random variables.

Definition 1. *Given a probability space $(\Omega, \mathbb{P}, \mathcal{F})$, a random variable $X : \Omega \mapsto \mathbb{R}$ is a function such that for any Borel set $\mathcal{B} \subset \mathbb{R}$, $X^{-1}(\mathcal{B}) \in \mathcal{F}$. Given a random variable X, the cumulative distribution function (cdf) of X is given by*

$$F_X(x) := P(X \leq x).$$

The cdf of a random variable essentially forgets the probability space but it bears all the interesting probabilistic properties of the given random variable. In some sense, this is the skeleton of the random variable. We say two random variables X and Y are equal in distribution if $F_X = F_Y$ and in that case we write $X \stackrel{d}{=} Y$. Note that, we do not require X and Y to be defined on the same probability space.

Another way to encode the probabilistic information of a random variable is to consider what is known as *characteristic function.*

Definition 2. *For a random variable X, the characteristic function of X, $\phi_X(t) : \mathbb{R} \to \mathbb{C}$ is given by*

$$\phi_X(t) := \mathbb{E}(e^{\mathbf{i}tX}) = \mathbb{E}(\cos(tX)) + \mathbf{i}\mathbb{E}(\sin(tx)).$$

The characteristic function uniquely determines the cdf of a random variable. It enjoys many interesting properties including uniform continuity. We refer to Theorem 3.3.1 and Theorem 3.3.2 in [1] for details. One of the interesting property of characteristic function is that is of interest to us is its connection to moments.

Theorem 2.1. *If $\mathbb{E}(X^n) < \infty$, then ϕ_X has continuous derivative of order n and*

$$\phi_X^{(n)}(t) = \mathbf{i}^n \mathbb{E}(X^n e^{\mathbf{i}tX}).$$

In particular, $\phi_X^{(n)}(0) = \mathbf{i}^n \mathbb{E}(X^n)$. Loosely speaking the sequence of moments (if exists) can be viewed as a sequence of derivatives of characteristic function evaluated at 0. This will be the key result when we look into the moment problem.

Next we consider a sequence of random variables and define the notion of *tightness* and *convergence in distribution.*

Definition 3. *X_n, X are random variables not necessarily definedd on the same probability space. We say X_n is tight if for all $\varepsilon > 0$, there is $M = M(\varepsilon) > 0$ such that for all $n \geq 1$,*

$$\mathbb{P}(|X_n| \geq M) \leq \varepsilon.$$

We say X_n converges in distribution to X if $F_{X_n}(x) \to F_X(x)$ for all continuity points of F. We write $X_n \xrightarrow{d} X$ in that case.

If X_n converges in distribution to X, one can show X_n is tight. On the other hand, if X_n is tight, it is not necessary that it should converge in distribution to some random variable. However, we can always find a subsequence X_{n_k} which converges in distribution. We refer to Section 3 in [1] for proof of these results. Note that this result is analagous to the concept of convergence and boundedness for real numbers sequence. Indeed, based on this fact, tightness and convergence in distribution can be thought of as a boundedness property and convergence property for random variables respectively.

3. Carleman's criterion

As mentioned in the introduction, two important questions related to the moment problem are existence and determinacy. In this section, we focus only on the determinacy.

Definition 4. *We say a random variable X is determinate if its moments sequence defined in* (1.1) *uniquely determines the cdf of X.*

In this section, we prove the Carleman condition, which gives a sufficient criterion for a random variable to be determinate. The Carleman condition follows from a more general result, the Denjoy-Carleman theorem, which we prove here. At this point, we recall Theorem **??** that connects moment sequence to the derivatives of characteristic function. Based on this identification, the question of detereminancy boils down to asking whether the derivative sequence (all evaluated at 0) uniquely determines the function. With this motivation, we define the concept of quasianalyticity below.

Definition 5. *Let $\mathcal{M} = (M_k)_{k\geq 0}$ be a sequence of positive numbers. The* **Carleman class** *$C\{\mathcal{M}\}$ consists of all $\phi \in C^\infty(\mathbb{R})$ such that, for some $C > 0$,*

$$\sup_{\xi\in\mathbb{R}} |\phi^{(k)}(\xi)| \leq C^{k+1} M_k.$$

We say a Carleman class is **quasianalytic** *if the map*

$$\phi \mapsto (\phi^{(k)}(0))_{k\geq 0}$$

is injective.

A function is then in a quasianalytic Carleman class if the function has all derivatives vanishing at a point and vanishes identically.

Definition 6. *We say a sequence (a_k) is* **log convex** *if*

$$a_k \leq \sqrt{a_{k-1}a_{k+1}}.$$

The Denjoy-Carleman theorem furnishes the necessary and sufficient conditions for a Carleman class to be quasianalytic.

Theorem 3.1 (Denjoy-Carleman Theorem)**.** *Let $\mathcal{M}$ be a log-convex sequence of positive numbers. The Carleman class $C\{\mathcal{M}\}$ is quasianalytic if and only if*

$$\sum_{k\geq 0} \frac{M_k}{M_{k+1}} = \infty. \tag{3.2}$$

Proof. Assume

$$\sum_{k\geq 0} \frac{M_k}{M_{k+1}} = \infty.$$

Let $\phi \in C^\infty$ with ϕ admitting the bounds

$$\sup_{\xi\in\mathbb{R}}|\phi^{(k)}(\xi)| \leq C^{k+1}M_k, \tag{3.3}$$

i.e., $\phi \in C\{\mathcal{M}\}$. For an integer $p \geq 0$, denote

$$\mathcal{B} = \{\xi \in \mathbb{R} : |\phi^{(k)}(\xi)| \leq C^{k+1}e^{k-p}M_k \text{ for } 0 \leq k < p\}.$$

Then $\mathbb{R} = \mathcal{B}_0 \supset \mathcal{B}_1 \supset \cdots$. If ϕ vanishes at 0 along with all of its derivatives, then $0 \in \bigcap_{p\geq 0}\mathcal{B}$. We want to show that if a point $\xi \in \mathcal{B}_p$ then there is $\delta > 0$ such that $(\xi-\delta, \xi+\delta) \in \mathcal{B}_{p-1}$. We do so by proving the contrapositive in the following lemma:

Lemma 3.2. *Assume that ϕ satisfies* (3.3) *for a log-convex sequence $\mathcal{M}$. If $\xi \notin \mathcal{B}_{p-1}$ for some $p \geq 1$, then*

$$\left[\xi - \frac{M_{p-1}}{CeM_p}, \xi + \frac{M_{p-1}}{CeM_p}\right] \cap \mathcal{B}_p = \emptyset.$$

Proof. Assume that $\xi + h \in \mathcal{B}_p$ for some $|h| \leq \frac{M_{p-1}}{CeM_p}$. Then for $k \leq p-1$

$$|\phi^{(k)}(\xi)| \leq \sum_{j=0}^{p-k-1}|\phi^{(k+j)}(\xi+h)\frac{|h|^j}{j!}| + C^{p+1}|\phi^{(p)}(\xi+\tilde{h})|\frac{|h|^{p-k}}{(p-k)!} \tag{3.4}$$

$$\leq \sum_{j=0}^{p-k}C^{k+j+1}e^{k+j-p}M_{k+j}\frac{|h|^j}{j!} \tag{3.5}$$

where we bounded the remainder by log-convexity and the definition of $\mathcal{B}_p$. By log-convexity

$$\begin{aligned}(3.5) = M_k\sum_{j=0}^{p-k}C^{k+j+1}e^{k+j-p}\frac{M_{k+j}}{M_k}\frac{|h|^j}{j!} &\leq M_k\sum_{j=0}^{p-k}C^{k+j+1}e^{k+j-p}\left(\frac{M_{k+j}}{M_k}\right)^j\frac{|h|^j}{j!}\\ &\leq M_k\sum_{j=0}^{\infty}C^{k+j+1}e^{k+j-p}\left(\frac{M_{k+j}}{M_k}\right)^j\frac{|h|^j}{j!} \leq M_kC^{k+1}e^{k-p+1}.\end{aligned}$$

□

If ϕ is not identically 0, there exists p and ξ such that $\xi \notin \mathcal{B}_p$. By the lemma, for $q > p$

$$\left[\xi - \sum_{k=p+1}^{q}\frac{M_{k-1}}{CeM_k}, \xi + \sum_{k=p+1}^{q}\frac{M_{k-1}}{CeM_k}\right] \cap \mathcal{B}_p = \emptyset.$$

Therefore as $\sum_{k=1}^\infty \frac{M_{k-1}}{M_k} = \infty$, we can find sufficiently large q such that $0 \notin \mathcal{B}_q$. Thus ϕ does not vanish at 0 with all derivatives. This is a contradiction. Therefore ϕ is identically 0 and so $C\{\mathcal{M}\}$ is quasianalytic.
We now prove necessity by assuming

$$\sum_{k=0}^{\infty}\frac{M_k}{M_{k+1}} < \infty$$

We construct a non-zero function with compact support $\phi \in C\{\mathcal{M}\}$ that is not quasianalytic. To see why this is sufficient, let $\phi \in C\{\mathcal{M}\}$ be such a function with compact support K. Then for some k, $\sup|\phi^{(k)}| \leq CM_k$. Then for all $x \in K$, it follows that $\sup|\phi^{(k-1)}| \leq CM_k|x|$. Likewise for $k-2$, $\sup|\phi^{(k-2)}| \leq CM_k\frac{|x|^2}{2}$. Then

$$\sup_{\xi}|\phi(\xi)| \leq CM_k\frac{|x|^k}{k!}.$$

therefore $\phi(0) = 0$, but $0 \in C\{\mathcal{M}\}$. Thus the map from ϕ to $(\phi^{(k}(0))_{k\geq 0}$ is not injective so ϕ cannot be quasianalytic.
Let u be a bump function such that $0 \leq u \leq 1$, $\;\operatorname{supp} u \subset [-2,2]$, and

$$\int u(\xi)d\xi = 1, \int |u'(\xi)|d\xi \leq 1.$$

Let $M_{-1} = \frac{M_0^2}{M_1}$. Define a sequence $u_k(\xi) = \frac{M_k}{M_{k-1}}u(\xi\frac{M_k}{M_{k-1}})$, and let

$$\phi_p = u_0 * u_1 * \cdots * u_{p-1}$$

be the convolution of the first p u_k. Then if $p > k$, ϕ_p admits the bounds

$$\begin{aligned}
|\phi_p^{(k)}| &= |u_0' * u_1' * \cdots * u_{k-1}' * u_k * \cdots u_{p-1}| \\
&\leq \prod_{j=0}^{k-1}\int |u_j'(\xi|d\xi(\max_{\xi}|u_k\xi|)\prod_{j=k+1}^{p-1}\int|u_j(\xi|d\xi \\
&\leq \prod_{j=0}^{k-1}\left(\frac{M_j}{M_{j-1}}\right)\left(\frac{M_k}{M-k-1}\right) = \frac{M_k}{M_{-1}}.
\end{aligned}$$

Then

$$\sup_{p>k}|\phi_p^{(k)}| \leq \frac{M_k}{M_{k-1}}$$

and so $\lim_{p\to\infty}\phi_p = \phi$ and the convergence is uniform. Thus we choose C sufficiently large so that $\phi \in C\{\mathcal{M}\}$. Then

$$\operatorname{supp}\,\phi \in \left[-2\sum_{k\geq 0}\frac{M_{k-1}}{M_k}, 2\sum_{k\geq 0}\frac{M_{k-1}}{M_k}\right].$$

So ϕ has compact support. $\square$

Corollary 3.3. *Let $\mathcal{M}$ be a sequence of positive numbers where $(M_k) \in \mathcal{M}$ is assumed log-convex. The Carleman class $C\{\mathcal{M}\}$ is quasianalytic if and only if*

$$\sum_{k\geq 1} M_k^{-1/k} = \infty, \tag{3.6}$$

then $C\{\mathcal{M}\}$ is quasianalytic.

Proof. We first assume quasianalyticity and show that (3.6) holds. Towards this end, we claim that log convexity implies

$$M_k \leq M_0\left(\frac{M_k}{M_{k-1}}\right)^k.$$

We proceed by induction. For $k = 2$, we note that $M_2 \leq M_0\left(\frac{M_2}{M_1}\right)^2$ is equivalent to

$$M_1 \leq \sqrt{M_0M_2}$$

which is log-convexity. Thus the claim holds for $k=2$. We assume it holds up to k. Note that by log-convexity we have

$$M_{k-1} \leq \sqrt{M_{k-2}M_k} \quad \text{and} \quad M_k \leq \sqrt{M_{k-1}M_{k+1}}$$

Taking product of the above two inequalities and rearranging terms we get

$$\frac{M_{k-1}}{M_{k-2}} \leq \frac{M_{k+1}}{M_k}.$$

Using log-convexity, induction hypothesis for $k-1$, and the above inequality we see that

$$M_{k+1} \leq M_{k-1}\left(\frac{M_{k+1}}{M_k}\right)^2 \leq M_0\left(\frac{M_{k-1}}{M_{k-2}}\right)^{k-1}\left(\frac{M_{k+1}}{M_k}\right)^2 \leq M_0\left(\frac{M_{k+1}}{M_k}\right)^{k+1}.$$

This proves the claim. Thus

$$\sum_{k\geq 1} M_k^{-1/k} \geq M_0^{-1/k}\sum_{k\geq 1}\frac{M_{k-1}}{M_k} = \infty$$

This proves (3.6). For the other direction we consider the following lemma.

Lemma 3.4 (Carleman's Inequality)**.** *For any positive sequence (a_k),*

$$\sum_{k=1}^{\infty}(a_1\cdot a_2\cdot \ldots \cdot a_k)^{1/k} \leq e\sum_{k=1}^{\infty} a_k$$

Proof. Let (r_k) be a sequence of positive real numbers.

$$(a_1 \ldots a_k)^{1/k} \leq \left(\frac{a_1r_1\ldots a_kr_k}{r_1\ldots r_k}\right)^{1/k} \leq \frac{a_1r_1+\ldots+a_kr_k}{k(r_1\ldots r_k)^{1/k}} \leq \sum_{j=1}^{\infty} a_jr_j\sum_{k\geq j}\frac{1}{k(r_1\ldots r_k)^{1/k}}$$

Therefore

$$\sum_{k=1}^{\infty}(a_1\cdot a_2\cdot \ldots \cdot a_k)^{1/k} \leq \sum_{k=1}^{\infty}\frac{a_1r_1+\ldots+a_kr_k}{k(r_1\ldots r_k)^{1/k}} \leq \sum_{j=1}^{\infty} a_jr_j\sum_{k\geq j}\frac{1}{k(r_1\ldots r_k)^{1/k}}.$$

Now choose $r_k = (k+1)^k/k^{k-1}$. Then $r_1\ldots r_k = (k+1)^k$. Then

$$\sum_{j=1}^{\infty} a_jr_j\sum_{k\geq j}\frac{1}{k(r_1\ldots r_k)^{1/k}} = \sum_{j=1}^{\infty} a_j\frac{(j+1)^j}{j^{j-1}}\sum_{k\geq j}\frac{1}{k(k+1)} = \sum_{j=1}^{\infty} a_j(1+1/j)^j$$
$$\leq e\sum_{j=1}^{\infty} a_j$$

where the last inequality follows from the limit representation of e. ☐

Turning back to the other direction of Corollary 3.3, we see that taking $a_k = \frac{M_k}{M_{k-1}}$ in Lemma 3.4, (3.6) implies (3.2). This completes the proof. ☐

From the Denjoy-Carleman Thoerem, we deduce the following criterion that gives us sufficiency for a measure to be determinate. This result will be crucial to our proof of the central limit theorem.

Corollary 3.5 (Carleman's Condition)**.** *Suppose X is a random variable*

$$\sum_{k=0}^{\infty} s_{2k}[\mu]^{-1/2k} = \infty$$

is determinate.

Proof. Let $\mathcal{M} = (M_k)_{k\geq 0}$, where $M_k = \sqrt{\mathbb{E}(X^{2k})}$. By Cauchy-Schwarz inequality, $\mathcal{M}$ is log-convex. So by the Denjoy-Carleman theorem and Corollary 3.3, $C\{\mathcal{M}\}$ is quasianalytic. We consider $\phi = \phi_X$, the characteristic function of the random variable X. Then

$$\sup_t |\phi_X(t)^k| = |\mathbb{E}(X^k e^{itX})| \leq \mathbb{E}|X^k| \leq \sqrt{\mathbb{E}(X^{2k})} = M_k.$$

Therefore $\phi_X \in C\{\mathcal{M}\}$. Thus the map is injective and so X is determinate. □

4. The Central Limit Theorem via Moments

Theorem 4.1 (Central Limit Theorem)**.** *Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. random variables with $\mathbb{E}(Z_1) = 0$ and $\mathbb{E}(Z_1^2) = 1$. Then*

$$X_n := \frac{Z_1 + Z_2 + \cdots + Z_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1). \tag{4.7}$$

Since our limiting random variable is Gaussian, it is useful to collect some of the properties of this distribution.

Proposition 4.2. *Let $X \sim \mathcal{N}(0, 1)$. Then for $k \geq 1$,*

$$\mathbb{E}(X^k) = \left\{ \begin{array}{ll} 0, & \textit{if } k \textit{ odd} \\ (k-1)(k-3)\cdot \cdots \cdot 5 \cdot 3 \cdot 1, & \textit{if } k \textit{ even} \end{array} \right\}$$

Furthermore, the moment sequence satisfies Carleman's condition and hence normal distribution is determinate.

Proof. We denote the pdf of the standard normal by $\phi(x)$. Recall that

$$\phi(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Let $k = 1$. Thus

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\phi(x)dx = \int_{-\infty}^{0} x\phi(x)dx + \int_{0}^{\infty} x\phi(x)dx.$$

Let $u = \frac{x^2}{2}$. Then by substitution

$$\int_{-\infty}^{0} x\phi(x)dx + \int_{0}^{\infty} x\phi(x)dx = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-z}dz + \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z}dz = -\frac{1}{\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}} = 0.$$

Note that $\phi'(x) = -x\phi(x)$. Then letting $u = x^{k-1}$ and $dv = \phi'(x)dx$ and integrating by parts

$$\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^{k-1}\phi'(x)dx = \int_{\infty}^{\infty} udv = x^{k-1}\phi(x)\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} (k-1)x^{k-2}\phi(x)dx = (k-1)\mathbb{E}(X^{k-2}).$$

Therefore $\mathbb{E}(X^2) = 1 \cdot \mathbb{E}(X^0) = \int_{-\infty}^{\infty} \phi(x)dx = 1$. Thus if k is even, the k-th moment is $(k-1)(k-3)\ldots \cdot 1$ and if k is odd the k-th moment is 0. One can check that it satisfies Carleman's condition using Stirling's approximation. □

Theorem 4.3. *Let $\{X_n\}$ be a sequence of random variables converging to a random variable X such that $\mathbb{E}(X_n^k) \to \mathbb{E}(X^k)$ (assume the expectations are finite) and let X be determinate. Then $X_n \xrightarrow{d} X$.*

Remark. The upshot of the above two results is that since Normal distribution is determinate, invoking Theorem 4.3 it is enough to show moment convergence to deduce convergence in distribution.

Proof. We first prove that $\{X_n\}$ is tight in order to find a subsequence X_{n_k} that converges to X in distribution. We then show this is independent of the subsequence.
By assumption

$$\mathbb{E}(X_n^2) \to \mathbb{E}(X^2)$$

so $\sup_n \mathbb{E}(X_n^2) < \infty$. By Markov's Inequality

$$\mathbb{P}(|X_n| > M) = \mathbb{P}(X_n^2 > M^2) \leq \frac{\mathbb{E}(X_n^2)}{M^2}.$$

Let $C = \sup_n \mathbb{E}(X_n^2)$. Fix $\epsilon > 0$. Then there is some $M > 0$ such that

$$\sup_n \mathbb{P}(X_n^2 > M^2) \leq \sup_n \frac{\mathbb{E}(X_n^2)}{M^2} = C/M^2 < \epsilon.$$

Thus $\{X_n\}$ is tight. Now let $\{X_{n_k}\}$ be an arbitrary subsequence of $\{X_n\}$ converging to a random variable Y. Recall that if h is a bounded, continuous function then for a sequence of random variables $\{Z_j\}$ converging to a random variable Z the following convergence holds:

$$\mathbb{E}(h(Z_j)) \to \mathbb{E}(h(Z)). \tag{4.8}$$

We let $h(x) = x^r$ with $r \geq 1$. We claim that

$$\mathbb{E}(X_{n_k}^r) \to \mathbb{E}(Y^r).$$

We will prove it for $r = 1$. The general case is analogous. Note that h is in fact unbounded. However, we can still make use of the result (4.8) by truncation. We impose the following constraint on h:

$$h_M(x) = \left\{ \begin{array}{ll} x, & \text{if } x \in [-M, M] \\ \pm M, & \text{if } x \notin [-M, M] \end{array} \right\}$$

Then

$$\begin{aligned} \mathbb{E}(X_{n_k}) - \mathbb{E}(Y) &= \mathbb{E}(X_{n_k}) - \mathbb{E}(h_M(X_{n_k})) + \mathbb{E}(h_M(X_{n_k})) - \mathbb{E}(h_M(Y)) + \mathbb{E}(h_M(Y)) - \mathbb{E}(Y) \\ &= \mathbb{E}(X_{n_k} - h_M(X_{n_k})) + \mathbb{E}(h_M(X_{n_k}) - h_M(Y)) + \mathbb{E}(h_M(Y) - Y) \end{aligned}$$

We show that each expectation above tends to 0 when we take $k \to \infty$ and then $M \to \infty$. From the definition of h_M,

$$\mathbb{E}(X_{n_k} - h_M(X_{n_k})) = \mathbb{E}(X_{n_k} \mathbb{1}_{X n_k \notin [-M,M]})$$

Then

$$\lim_{M\to\infty} \mathbb{E}(X_{n_k} \mathbb{1}_{X_{n_k} \notin [-M,M]}) \leq \lim_{M\to\infty} \sup_n \mathbb{E}(X_{n_k} \mathbb{1}_{X_{n_k} \notin [-M,M]}) = 0.$$

$$\begin{aligned} |\mathbb{E}(h_M(Y) - Y)| &\leq \mathbb{E}(|Y - M| \mathbb{1}_{Y \notin [-M,M]}) \\ &\leq \mathbb{E}(|Y| \mathbb{1}_{Y \notin [-M,M]}) + \mathbb{E}(|M| \mathbb{1}_{Y \notin [-M,M]}). \end{aligned}$$

Define a sequence Y_M so that $Y_M \leq |Y|$ by

$$Y_M = \mathbb{E}(|Y| \mathbb{1}_{Y \notin [-M,M]}).$$

Clearly $Y_M \to 0$ pointwise as $M \to \infty$. As $|Y_M| \leq |Y|$ and $\mathbb{E}(Y)$ exists and is finite, $\mathbb{E}(Y_M) \to 0$ as $M \to \infty$ by the dominated convergence theorem. For the second half we make use of Markov's inequality:

$$\mathbb{E}(|M| \mathbb{1}_{Y \notin [-M,M]}) \leq |M| \mathbb{P}(Y \notin [-M, M]) = \mathbb{P}(|Y| > M) \leq |M| \frac{\mathbb{E}(Y)}{M^2} = C/M.$$

C/M tends to 0 as $M \to \infty$. And finally using uniform integrability one can show under the double limit

$$|\mathbb{E}(h_M(X_{n_k})) - \mathbb{E}(h_M(Y))| \to 0$$

We omit the calculations here. Therefore

$$\limsup_{n\to\infty} |\mathbb{E}(X_{n_k}) - \mathbb{E}(Y)| \le \lim_{M\to\infty} \sup_n |\mathbb{E}(X_{n_k} - h_M(X_{n_k}))| + \lim_{n\to\infty} |\mathbb{E}(h_M(X_{n_k}) - h_M(Y))| \\ + \lim_{M\to\infty} |\mathbb{E}(h_M(Y) - Y)| \to 0.$$

Therefore $E(X_{n_k}) \to \mathbb{E}(Y)$. As mentioned before, similar arguments lead to $\mathbb{E}(X_{n_k}^r) \to \mathbb{E}(Y^r)$ for all $r \ge 1$. But X is determinate, so then $X \stackrel{d}{=} Y$. Thus $X_{n_k} \underset{d}{\to} X$. Thus every subsequence of X_n has a further subsequence which converges in distribution to X. Hence $X_n \xrightarrow{d} X$. □

4.1. **Proof of Theorem 4.1.** The proof of Theorem 4.1 is carried in following 4 steps.

Step 1. Truncation. We wish to compute the moments of X_n defined in (4.7). However, the higher moments of X_n may not even exist. One way to get rid of this hurdle is to consider *truncated random variables.* Fix $M > 0$. Define $Y_{i,n} = Z_i \mathbb{1}_{|Z_i| \le M\sqrt{n}}$ and then define

$$W_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_{i,n}.$$

The advantage of this truncation is that since $|Y_{i,n}| \le M\sqrt{n}$, we have $|W_n| \le Mn$, so $\mathbb{E}(|W_n|^k) < \infty$ for each k and for each n. So, the moment sequence is well defined. However, the problem is $\mathbb{E}(Y_{i,n})$ may not be zero and variance may not be 1. So, we rescale and recenter! Define

$$U_{i,n} := \frac{Y_{i,n} - \mathbb{E}(Y_{i,n})}{\sqrt{\mathrm{Var}(Y_{i,n})}}, \qquad V_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n U_{i,n}.$$

Our first step is to show that

$$\mathbb{E}(V_n^k) \xrightarrow{d} \mathcal{N}(0,1).$$

To prove this, we use Carleman's condition and Theorem 4.3. We specifically rely on a truncation scheme as in Theorem 4.3 so we must verify that $\mathbb{E}(V_n^k)$ is bounded for all k.
We begin by making sure $\mathbb{E}(Y_{i,n})$ is bounded and that $\mathrm{Var}(Y_{i,n})$ is bounded and not zero. Note that $Y_{i,n}$ is independent of i, so we take $i = 1$ without loss of generality.
Observe that Z_1 can be expressed in the following way

$$Z_1 = Z_1 \mathbb{1}_{|Z_1| \le M\sqrt{n}} + Z_1 \mathbb{1}_{|Z_1| > M\sqrt{n}}.$$

By Markov's Inequality

$$\mathbb{P}(|Z_1| > \sqrt{n}M) = M^2 n \mathbb{E}(\mathbb{1}_{|Z_1| > \sqrt{n}M}) \le \mathbb{E}(Z_1^2)$$

Observe that

$$M^2 n \mathbb{1}_{|Z_1| > M\sqrt{n}} \le M\sqrt{n}|Z_1| \le M\sqrt{n} Z_1^2.$$

The above inequalities imply

$$\mathbb{E}(\mathbb{1}_{|Z_1| > \sqrt{n}M}) \le \frac{\mathbb{E}(|Z_1|)}{M\sqrt{n}} \le \frac{\mathbb{E}(Z^2)}{M\sqrt{n}} = \frac{1}{\sqrt{n}M}.$$

Thus $\mathbb{P}(|Z_1| \ge \sqrt{n}M) \to 0$ as $n \to \infty$. Note that $Y_{1,n}$ can be expressed in the following way

$$Y_{1,n} = Z_1 \mathbb{1}_{|Z_1| \le M\sqrt{n}} = Z_1 - Z_1 \mathbb{1}_{|Z_1| > M\sqrt{n}}$$

so $|Y_{1,n}| \leq Z_1$. Thus

$$\lim_{n\to\infty} \mathbb{E}(|Y_{1,n}|) = \lim_{n\to\infty} \mathbb{E}(|Z_1 - Z_1 \mathbb{1}_{|Z_1|)>M\sqrt{n}}|) \leq \mathbb{E}(Z_1) + \lim_{n\to\infty} \mathbb{E}(Z_1 \mathbb{1}_{|Z_1|>M\sqrt{n}}) \leq \mathbb{E}(Z_1) + \frac{1}{M\sqrt{n}}$$

so by the dominated convergence theorem as $n \to \infty$

$$\mathbb{E}(Y_{1,n}) \to 0 \quad \text{which implies} \quad \mathbb{E}(Y_{1,n}) \to \mathbb{E}(Z_{1,n}).$$

We now consider $\mathrm{Var}(Y_{1,n})$. Using our results from above,

$$\mathrm{Var}(Y_{1,n}) = \mathbb{E}(Y_{1,n}^2) - \mathbb{E}(Y_{1,n})^2 \leq \mathbb{E}(Y_{1,n}^2).$$

By expresssing $\mathbb{E}(Y_{1,n}^2)$ as

$$\mathbb{E}(Y_{1,n}^2) = \mathbb{E}(\left[Z_1 - Z_1 \mathbb{1}_{|Z_1|>M\sqrt{n}}\right]^2)$$

we see that

$$\begin{aligned} \mathbb{E}(Y_{1,n}^2) &= \mathbb{E}(Z_1^2 - 2Z_1^2 \mathbb{1}_{|Z_1|>M\sqrt{n}} + Z_1^2 \mathbb{1}_{|Z_1|>M\sqrt{n}}) \\ &\leq \mathbb{E}(Z_1^2) + \mathbb{E}(Z_1^2 \mathbb{1}_{|Z_1|>M\sqrt{n}}) = \mathbb{E}(Z_1^2) + \frac{1}{M^2 n} \end{aligned}$$

Applying the dominated convergence theorem again, we see that

$$\mathbb{E}(Y_{1,n}^2) \to \mathbb{E}(Z_1^2)$$

so

$$\mathrm{Var}(|Y_{1,n}|) \leq \mathbb{E}(Y_{1,n}^2) + \left(\frac{1}{M\sqrt{n}}\right)^2 = 1 + \frac{2}{M^2 n} \to E(Z_1^2) = \mathrm{Var}(Z_1) = 1$$

as $n \to \infty$.
We now need to show the moments of V_n are bounded, i.e., for all k, $\mathbb{E}(V_n^k) < \infty$. Note that

$$V_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{Y_{i,n} - \mathbb{E}(Y_{i,n})}{\sqrt{\mathrm{Var}(Y_{i,n})}} = \frac{n}{\sqrt{n}} \left(\frac{Y_{1,n}}{\sqrt{\mathrm{Var}(Y_{1,n})}} - \frac{\mathbb{E}(Y_{1,n})}{\sqrt{\mathrm{Var}(Y_{i,n})}} \right) = \frac{\sqrt{n} Y_{i,n}}{\sqrt{\mathrm{Var}(Y_{i,n})}} - \frac{\sqrt{n}\mathbb{E}(Y_{i,n})}{\sqrt{\mathrm{Var}(Y_{i,n})}}.$$

Then

$$\begin{aligned} \mathbb{E}(|V_n|) &= \mathbb{E}\left(\left| \frac{\sqrt{n} Y_{1,n}}{\sqrt{\mathrm{Var}(Y_{1,n})}} - \frac{\sqrt{n}\mathbb{E}(Y_{1,n})}{\sqrt{\mathrm{Var}(Y_{1,n})}} \right| \right) \leq \mathbb{E}\left(\left| \frac{\sqrt{n} Y_{1,n}}{\sqrt{\mathrm{Var}(Y_{1,n})}} \right| \right) + \mathbb{E}\left(\left| \frac{\sqrt{n}\mathbb{E}(Y_{1,n})}{\sqrt{\mathrm{Var}(Y_{1,n})}} \right| \right) \\ &\leq \left| \frac{\frac{1}{M}}{\sqrt{1 + \frac{1}{M^2 n}}} \right| + \left| \frac{\frac{1}{M}}{\sqrt{1 + \frac{1}{M^2 n}}} \right| = \frac{2}{\sqrt{M^2 + \frac{1}{n}}}. \end{aligned}$$

From the case $k = 1$, it follows that for $k > 1$

$$\mathbb{E}(|V_n|^k) \leq \mathbb{E}\left(\left| \left[\frac{2}{\sqrt{M^2 + \frac{1}{n}}} \right] \right|^k \right) = \frac{2^k}{(M^2 + \frac{1}{n})^{k/2}}.$$

As $\mathbb{E}(V_n^k) \leq \mathbb{E}(|V_n|^k)$ for all k, the moments of V_n are bounded. Next we compute asymtptics of moments of V_n. We claim that

$$\mathbb{E}(V_n^{2k}) \to 1 \cdot 3 \cdot 5 \cdots (2k-1) \quad \text{and} \quad \mathbb{E}(V_n^{2k-1}) \to 0 \tag{4.9}$$

for $k \geq 1$. We first note that by assumption

$$\mathbb{E}(V_n) \to 0 \text{ and } \mathbb{E}(V_n^2) \to 1$$

Step 2: $k = 3$ and $k = 4$. We establish the moments of $k = 3$ and $k = 4$.
For V_n^3 we find the following

$$\mathbb{E}(V_n^3) = \mathbb{E}\left(\frac{[U_1 + ... + U_n]^3}{n^{3/2}}\right) = \mathbb{E}\left(\sum_i^n \frac{U_i^3}{n^{3/2}} + 3\sum_{i\neq j} \frac{U_i^2 U_j}{n^{3/2}} + \sum_{i\neq j\neq r} \frac{U_i U_j U_r}{n^{3/2}}\right). \tag{4.10}$$

Note that if $i \neq j$ U_i and U_j are independent of one another. Then $\mathbb{E}(U_i U_j) = \mathbb{E}(U_i)\mathbb{E}(U_j)$. By independence and linearity of expectation

$$(4.10) = \sum_{i=1}^n \frac{\mathbb{E}(U_i^3)}{n^{3/2}} + 3\sum_{i\neq j}^n \frac{\mathbb{E}(U_i^2)\mathbb{E}(U_j)}{n^{3/2}} + \sum_{i\neq j\neq r}^n \frac{\mathbb{E}(U_i)\mathbb{E}(U_j)\mathbb{E}(U_r)}{n^{3/2}} \leq \frac{nC_3}{n^{3/2}}$$

where C_3 is the value of $\mathbb{E}(U_i^3)$. Clearly

$$\frac{nC_3}{n^{3/2}} = \frac{C_3}{\sqrt{n}} \to 0 \quad \text{as} \quad n \to \infty$$

so $\mathbb{E}(V_n^3) \to 0$. We now consider $\mathbb{E}(V_n^4)$:

$$\begin{aligned}\mathbb{E}(V_n^4) = \mathbb{E}\left(\frac{(U_1 + ... + U_n)^4}{n^2}\right) &= \frac{\sum_i \mathbb{E}(U_i^4)}{n^2} + 4\frac{\sum_{i\neq j} \mathbb{E}(U_i^3 U_j)}{n^2} \\ &+ 3\frac{\sum_{i\neq j} \mathbb{E}(U_i^2 U_j^2)}{n^2} + 6\frac{\sum_{i\neq j\neq r} \mathbb{E}(U_i^2 U_j U_r)}{n^2} + \frac{\sum_{i\neq j\neq r\neq q} \mathbb{E}(U_i U_j U_r U_q)}{n^2} \\ &= \frac{\sum_i \mathbb{E}(U_i^4)}{n^2} + 4\frac{\sum_{i\neq j} \mathbb{E}(U_i^3)\mathbb{E}(U_j)}{n^2} + 3\frac{\sum_{i\neq j} \mathbb{E}(U_i^2)\mathbb{E}(U_j^2)}{n^2} \\ &+ 6\frac{\sum_{i\neq j\neq r} \mathbb{E}(U_i^2)\mathbb{E}(U_j)\mathbb{E}(U_r)}{n^2} + \frac{\sum_{i\neq j\neq r\neq q} \mathbb{E}(U_i)\mathbb{E}(U_j)\mathbb{E}(U_r)\mathbb{E}(U_q)}{n^2}\end{aligned}$$

From the argument for the previous case, all summands with a quantity to the first power go to 0. All that remains then is the following two terms:

$$\frac{\sum_i \mathbb{E}(U_i^4)}{n^2} + 3\frac{\sum_{i\neq j} \mathbb{E}(U_i^2)\mathbb{E}(U_j^2)}{n^2}.$$

From the fact that $\mathbb{E}(U_i^2) \to 1$ and that there are $n(n-1)$ combinations of i, j withthe condition $i \neq j$ the above equation simplifies to

$$\frac{nC_4}{n^2} + 3\frac{n(n-1)}{n^2}$$

where $C_4 = \mathbb{E}(U_i^4)$. Then

$$\mathbb{E}(V_n^4) = \frac{C_4}{n} - \frac{3}{n} + 3 = O(1/n) + 3 \to 3.$$

Step 3: General moments. We now prove the general case. We claim that all terms in the expansion of $\mathbb{E}\left(\frac{(U_1+U_2+...U_n)^t}{n^{t/2}}\right)$ tend asymptotically to zero with exception of the term where each U_i is to the second power. Then if $t = 2k - 1$, $\mathbb{E}(V_n^t) \to 0$ and if $t = 2k$ then $\mathbb{E}(V_n^t) \to (2k - 1) \cdot ... \cdot 5 \cdot 3 \cdot 1$.
For arbitrary t

$$V_n^t = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^n U_i^t + ... + \sum_{i_1\neq i_2...\neq i_t}^n U_{i_1} U_{i_2} ... U_{i_t}\right)$$

which we refer to as the expansion of V_n^t. Each summand is a partition of t, i.e., in any summand

$$\sum_{i_1 \neq i_2 \neq \ldots \neq i_l} U_{i_1}^{b_1} U_{i_2}^{b_2} \ldots U_{i_l}^{b_l}$$

the multiplicities $b_1, \ldots, b_l$ must sum to t. As $\mathbb{E}(V_n) \to 0$, the presence of any U_i with multiplicity 1 in a summand implies that the summand goes to 0. Then we can restrict ourselves to the case where $b_j > 1$ for all $1 \leq j \leq l$. We note then that the maximal number of multiplicities is $t/2$ if $t = 2k$ or $(t-1)/2$ if $t = 2k-1$ (if greater $t/2$, then the product must contain a term of multiplicity one, which we have disregarded). The the value of each summand if $t = 2k$ is bounded above by

$$n^{t/2} \frac{C_{b_1} C_{b_2} \ldots C_{b_{t/2}}}{n^{t/2}} = C_{b_1} C_{b_2} \ldots C_{b_k}.$$

If the number of multiplicities l is less than $t/2$, then the term has value

$$n(n-1)\ldots(n-l)\left(\frac{C_{b_1} C_{b_2} \ldots C_{b_l}}{n^{t/2}}\right).$$

But

$$n(n-1)\ldots(n-l)\left(\frac{C_{b_1} C_{b_2} \ldots C_{b_l}}{n^{t/2}}\right) \leq n^l \frac{C_{b_1} C_{b_2} \ldots C_{b_l}}{n^{t/2}} \leq \frac{C_{b_1} C_{b_2} \ldots C_{b_l}}{n} \to 0$$

as $n \to \infty$. Then if $l \neq t/2$, the term must tend to 0. If $l = t/2$, we have

$$n(n-1)\ldots(n-k+1)\frac{C_{b_1} C_{b_2} \ldots C_{b_k}}{n^k} = \left[\frac{n(n-1)\ldots(n-k)}{n^k}\right] C_{b_1} C_{b_2} \ldots C_{b_k}.$$

The quantity $\frac{n(n-1)\ldots(n-k)+1}{n^k} \to 1$ as $n \to \infty$ so as all multiplicities are two and $\mathbb{E}(V_n^2) \to 1$ so all $C_i = C_2$ so the term tends to 1.

Note that for the term with all multiplicities of 2, the quantity $(U_1 + \ldots + U_n)^{2k}$ has $2k-1$ choices of U_1 to make U_1^2, $2k-3$ choices for U_2, down to 1 choices for U_n. Putting this together with the fact the rest of the term tends to 1, we see that

$$\begin{aligned} &\frac{1}{n^k}(2k-1)(2k-3)\ldots \cdot 5 \cdot 3 \cdot 1 \sum_{i=1}^{n} \mathbb{E}\left(U_1^2 U_2^2 \ldots U_n^2\right) \\ &= (2k-1)(2k-3)\ldots 3 \cdot 1 \left[\frac{n(n-1)\ldots(n-k)}{n^k}\right] \to (2k-1)(2k-3)\ldots 3 \cdot 1 \end{aligned}$$

as $n \to \infty$.This is the only term in $(V_n)^k$ that does not asymptotically tend to 0, so

$$\mathbb{E}(V_n^{2k}) \to (2k-1)\ldots 5 \cdot 3 \cdot 1 \quad \text{as} \quad n \to \infty.$$

If $t = 2k-1$, there is no term consisting of all multiplicities 2, so each term in the expansion either tends to 0 from singletons or tends to 0 for large n. Thus

$$\mathbb{E}(V_n^{2k+1}) \to 0 \quad \text{as} \quad n \to \infty.$$

As we have explicitly established the moments for $k = 1, 2, 3, 4$, inductively our results for $2k-1$ and $2k$ must indeed hold in general. This completes the proof of (4.9). Invoking Proposition 4.2 and Theorem 4.3 we get that

$$V_n \xrightarrow{d} \mathcal{N}(0,1).$$

Step 4: Final Steps. We now show that $X_n \xrightarrow{d} \mathcal{N}(0,1)$, which will conclude our proof of the central limit theorem. As $\mathbb{E}(X_n^k)$ may not exist, we proceed by comparison of V_n and X_n. Our next step is to obtain inequalities in order to compare V_n and X_n. First let us compare V_n and W_n. Note that

$$W_n = \sqrt{\mathrm{Var}(Y_{1,n})}V_n + \sqrt{n}\mathbb{E}(Y_{1,n}).$$

Define $a_n = \sqrt{\mathrm{Var}(Y_{1,n})} \to 1$ and $b_n = \sqrt{n}\mathbb{E}(Y_{1,n})$. Note that

$$\mathbb{E}(Y_{1,n}) = \mathbb{E}(Z_1) - \mathbb{E}(Z_1 \mathbb{1}_{|Z_1|>M\sqrt{n}}) = -\mathbb{E}(Z_1 \mathbb{1}_{|Z_1|>M\sqrt{n}}).$$

Thus using the fact that $\mathbb{1}_{|Z_1|>M\sqrt{n}} \leq \frac{|Z_1|}{M\sqrt{n}}$ we get

$$|\sqrt{n}\mathbb{E}(Y_{1,n})| \leq \sqrt{n}\mathbb{E}(|Z_1|\mathbb{1}_{|Z_1|>M\sqrt{n}}) \leq \frac{1}{M}\mathbb{E}(Z_1^2) = \frac{1}{M}.$$

Thus $|b_n| \leq \frac{1}{M}$. Now let us compare W_n and X_n. Note that the event that $(W_n \neq X_n)$ is a subset of $\bigcup\limits_{i=1}^{n}(|Z_i| > M\sqrt{n})$. Thus

$$P(W_n \neq X_n) \leq \sum_{i=1}^{n} P(|Z_i| > M\sqrt{n}) \leq n\frac{1}{M^2 n}\mathbb{E}(Z_1^2) = \frac{1}{M^2}.$$

The event $(W_n = X_n)$ is a subset of $\bigcup\limits_{i=1}^{n}(|Z_i| \leq M\sqrt{n})$, so

$$\mathbb{P}(W_n = X_n) \leq \sum_{i=1}^{n} \mathbb{P}(\sqrt{n}|Z_i| \leq M\sqrt{n}) \leq \frac{1}{M}$$

by the argument for V_n and W_n. Therefore

$$\mathbb{P}(X_n \leq x) = \mathbb{P}(X_n \leq x, X_n = W_n) + \mathbb{P}(X_n \leq x, X_n \neq W_n) \leq \mathbb{P}(W_n \leq x) + \mathbb{P}(W_n \neq X_n).$$

Thus

$$\mathbb{P}(X_n \leq x) \leq \mathbb{P}(W_n \leq x) + \mathbb{P}(W_n \neq X_n) \tag{4.11}$$

and

$$\mathbb{P}(X_n \leq x) \geq \mathbb{P}(W_n \leq x) - \mathbb{P}(W_n \neq X_n). \tag{4.12}$$

Our final step is to complete the proof from definition. Consider n large enough such that $|a_n - 1| \leq \frac{1}{M}$. This restriction means

$$1 - \frac{1}{M} \leq a_n \leq 1 + \frac{1}{M}.$$

Fix $x \in \mathbb{R}$. For the case of (4.11) we take $-\frac{1}{M} \leq b_n \leq 0$. Then

$$\begin{aligned}
\mathbb{P}(X_n \leq x) &\leq \mathbb{P}(X_n \leq x, W_n = X_n) + \mathbb{P}(W_n \neq X_n) \\
&\leq \mathbb{P}(W_n \leq x) + \frac{1}{M^2} \\
&\leq \mathbb{P}(a_n V_n + b_n \leq x) + \frac{1}{M^2} \\
&\leq \mathbb{P}\left(V_n \leq \frac{x - b_n}{a_n}\right) + \frac{1}{M^2} \leq \mathbb{P}\left(V_n \leq \frac{x + \frac{1}{M}}{1 - \frac{1}{M}}\right) + \frac{1}{M^2}
\end{aligned}$$

But we have already established normal convergence for V_n. So the last thing converges to

$$\mathbb{P}\left(N(0,1) \leq \frac{x + \frac{1}{M}}{1 - \frac{1}{M}}\right) + \frac{1}{M^2}$$

Now we finally take M to ∞ to get that $\mathbb{P}(X_n \le x) \le \mathbb{P}(N(0,1) \le x)$.
For (4.12) we take $0 \le b_n \le \frac{1}{M}$. Then

$$\mathbb{P}(W_n \le x) = \mathbb{P}(W_n \le x, W_n = X_n) + \mathbb{P}(W_n \le x, W_n \ne X_n) \le \mathbb{P}(X_n) + \mathbb{P}(W_n \ne X)$$

$$\implies \mathbb{P}(X_n \le x) \ge \mathbb{P}(W_n \le x) - \mathbb{P}(W_n \ne X_n) = \mathbb{P}\left(V_n \le \frac{x - \frac{1}{M}}{1 + \frac{1}{M}}\right) - \frac{1}{M^2}.$$

The last quantity converges to

$$\mathbb{P}\left(V_n \le \frac{x - \frac{1}{M}}{1 + \frac{1}{M}}\right) - \frac{1}{M^2} \to \mathbb{P}\left(\mathcal{N}(0,1) \le \frac{x - \frac{1}{M}}{1 + \frac{1}{M}}\right) - \frac{1}{M^2}.$$

Then taking $M \to \infty$

$$\mathbb{P}(\mathcal{N}(0,1) \le x) \le \mathbb{P}(X_n \le x)$$

Putting the two cases together we now have that $\mathbb{P}(X_n \le x) \to \mathbb{P}(\mathcal{N}(0,1) \le x)$. Then $X_n \xrightarrow{d} \mathcal{N}(0,1)$. This completes the proof of the Central Limit Theorem.

Acknowledgement. The author thanks Sayan Das for mentoring him during the project. This reading project was carried out under Twoples mentoring program. Both mentor and mentee are very grateful to the organizer Stephen McKean for the matchmaking.

References

[1] Rick Durrett. *Probability: theory and examples*. Cambridge University Press, 2019.

University of California, Berkeley