# Central Limit Theorem

## Motivations, Proof, Applications, and More

Luke Bula

Euler Circle

July 12, 2026

# 1 A Beautiful Theorem

## According to Mathematicians

- Wendel H. Fleming: "**beautiful** *mathematical result*" and "*profoundly important for applications.*"
- Shige Peng: claim about the "**beauty** *and power*" of the result.
- Francis Galton: "*an unsuspected and most* **beautiful** *form of regularity.*"
- Thomas Sargent: "*one of the most* **remarkable** *results in all of mathematics.*"
- Alan Hajek: reference to "**beautiful** *theorems.*"

## 2 The Central Limit Theorem

**Theorem**

*The Central Limit Theorem Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$, where $0 < \sigma^2 < \infty$. Then*

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

# 3 Coin Flip Demonstration

- $Coin_1, Coin_2, \ldots, Coin_n$ are independentally and indentically distributed.
- $P(Coin_i = 1) = .5, P(Coin_i = 0) = .5$ for all $i$, $1 \leq i \leq n$

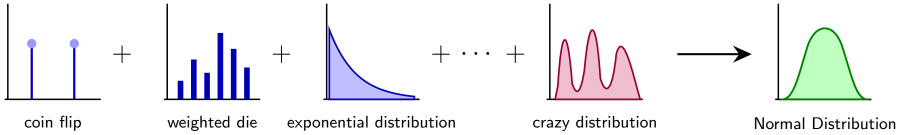
Simulation

# 4 Power of the Central Limit Theorem

## Too Perfect? Nothing Special?

- 50/50 Coin Flips
- Too Simple
- Identical Distributions

Any distribution works!

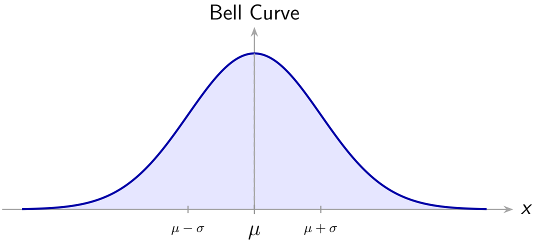



## Takeaways

All roads lead to normal!

# 5 What Is the Bell Curve?

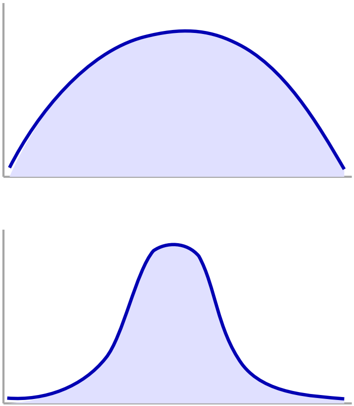
- Density of a normal, de Moivre's, or Gaussian distribution.
- $\mu$ shifts the center; $\sigma$ controls the spread.
- The standard normal $Z \sim \mathcal{N}(0,1)$ is the reference version.




**Try it yourself!**

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

# 6 Why Prove It?



# 7 If ..., then...

Core Idea: $X_1 + \cdots + X_n \to \mathcal{N}(n\mu, \sigma\sqrt{n})$ as $n \to \infty$

---

## Ways to Prove it

- Characteristic Functions
- Fourier Analysis
- Martingales
- Entropy/Information Theory
- Renormalization Group Theory
- Double Factorials and Combinatorics

We will use the **Moment Method**

# 8 Background

## Probability Density Functions

- A PDF $f_X(x)$ describes the shape of a continuous random variable (Probability Mass Functions work for discrete variables).
- Probabilities are areas under the curve:

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)\, dx.$$

- A PDF is unique to its corresponding distribution
- Hard to determine the PDF

## Moment Generating Functions

- An MGF turns a distribution into a function:

$$M_X(t) = \mathbb{E}\left[e^{tX}\right].$$

- Derivatives at 0 give moments: $M_X'(0) = \mathbb{E}[X]$.
- Moments are the fingerprints of the distribution.
- $M_{\mathcal{N}}(t) = e^{t^2/2}$
- Has useful properties for proving the CLT.

# 9 Our Lucky Lemma

Let $Z_1, Z_2, \ldots$ be random variables with PDFs $F_{Z_n}$ and moment generating functions $M_{Z_n}$. Let $Z$ have PDF $F_Z$ and moment generating function $M_Z$. If

$$M_{Z_n}(t) \to M_Z(t)$$

for all $t$, then

$$F_{Z_n}(x) \to F_Z(x)$$

for every point $x$ where $F_Z$ is continuous.

# 10 In Other Words...

Suppose $S_1, S_2, \ldots$, where $S_i = X_1 + \cdots + X_i$, are i.i.d. random variables with PDFs $F_{S_n}$ and moment generating functions $M_{S_n}$, and a normal distribution $\mathcal{N}$ has PDF $F_{\mathcal{N}}$ and moment generating function $M_{\mathcal{N}}$. If

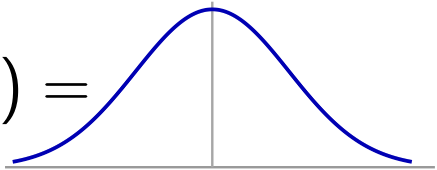
$$M_{S_n}(t) \to M_{\mathcal{N}}(t)$$

for all $t$, then

$$F_{S_n}(x) \to F_{\mathcal{N}}(x)$$

for every point $x$ where $F_Z$ is continuous.

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



# 11 Showing Convergence of MGFs

Assumptions: $\mu = 0, \sigma^2 = 1, i.i.d.$
Goal: $M_{\frac{S_n(t)}{\sqrt{n}}}(t) \to M_{\mathcal{N}(0,1)}(t)$

---

*Proof*
By the definition of the MGF, we see that the MGF of $X_i$, for some $i \in \mathbb{N}$, is given by

$$M_{X_i}(t) = \mathbb{E}\left[e^{tX_i}\right].$$

**Proposition** If some random variable $Y$ has a corresponding MGF given by $M_Y(t)$, then $M_{\frac{Y}{b}} = M_Y(\frac{t}{b})$ where $b \in \mathbb{R}$.
Thus, $M_{\frac{X_i}{\sqrt{n}}}(t) = M_{X_i}(\frac{t}{\sqrt{n}})$

**Proposition** If $X$ and $Y$ are independent random variables, then

$$M_{X+Y}(t) = M_X(t)\, M_Y(t).$$

Thus,

$$M_{\frac{S_n}{\sqrt{n}}}(t) = M_{X_1+\cdots+\ X_n}(\frac{t}{\sqrt{n}}) = \left[M_{X_i}(\frac{t}{\sqrt{n}})\right]^n.$$

# 11 Showing Convergence of MGFs

Proceeding, we will show

$$\lim_{n\to\infty}\left[M_{X_i}(\frac{t}{\sqrt{n}})\right]^n = e^{\frac{t^2}{2}}$$

Taking the natural log of both sides we obtain

$$\lim_{n\to\infty} nL(\frac{t}{\sqrt{n}}) = \frac{t^2}{2}.$$

where $L(t) = \log\left(M_{X^i}(t)\right)$.
Applying L'Hopital's Rule twice, we are left with

$$= \lim_{n\to\infty}\left[L''\left(\frac{t}{\sqrt{n}}\right)\frac{t^2}{2}\right],$$

which evaluates to

$$= \frac{t^2}{2}.$$

# How is it Useful?

It allows us to use the Gaussian to approximate sums or means of distributions. (n=30)

$$X_1 + \cdots + X_n \approx \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

# Example 1: Sum of 30 Dice Rolls

Let

$$S_{30} = X_1 + X_2 + \cdots + X_{30},$$

where each $X_i$ is one fair die roll.

$$\mu = 3.5, \qquad \sigma^2 = \frac{35}{12}$$

Therefore,

$$E[S_{30}] = 30(3.5) = 105, \qquad \mathrm{SD}(S_{30}) = \sqrt{30 \cdot \frac{35}{12}} \approx 9.35.$$

## Question

What is the probability that the sum is greater than 120?

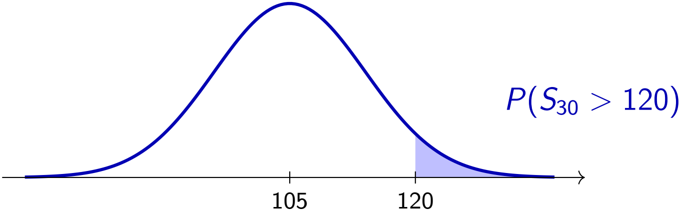
$$P(S_{30} > 120)$$

# Using the Central Limit Theorem

By the Central Limit Theorem,

$$S_{30} \approx N(105, 87.5).$$

Our Question Visually:




## Approximation

$$P(S_{30} > 120) \approx 0.049.$$

$\boxed{P(S_{30} > 120) \approx 4.9\%, .0003\% \text{ off exact answer}}$

# 12 Higher-Dimensional Central Limit Theorem

The Central Limit Theorem can also be generalized to random vectors.

Instead of one random variable,

$$X_i \in \mathbb{R},$$

we can study vectors,

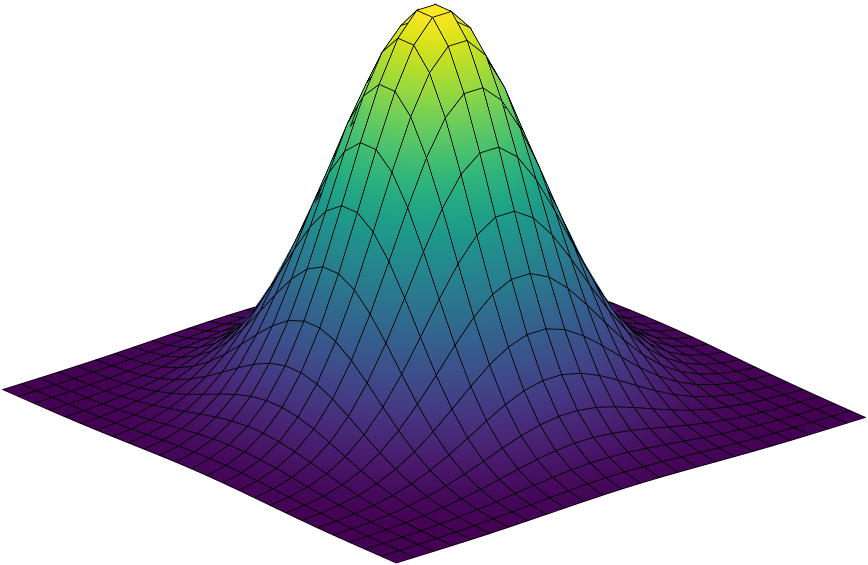
$$\mathbf{X}_i \in \mathbb{R}^d.$$

If the vectors are independent and have mean vector $\boldsymbol{\mu}$ and covariance matrix $\Sigma$, then

$$\sqrt{n}\left(\overline{\mathbf{X}}_n - \boldsymbol{\mu}\right) \Longrightarrow N(\mathbf{0}, \Sigma).$$

## Why this is useful

- modeling several related measurements at once
- estimating multiple averages from the same sample
- applications in physics, finance, biology, and data science



# The End

If you liked my talk, then read my paper!

- Everything Proved
- Bell Curve Intuition
- Exceptions to CLT
- Outline Fascinating Double Factorial Proof
- Connections to Physics
- Information Theoretic Lens
- Interesting Applications

Email: Lukejb0001@gmail.com