DCP 1206: Probability
Lecture 24 — Strong Law of Large Numbers and Central Limit Theorem

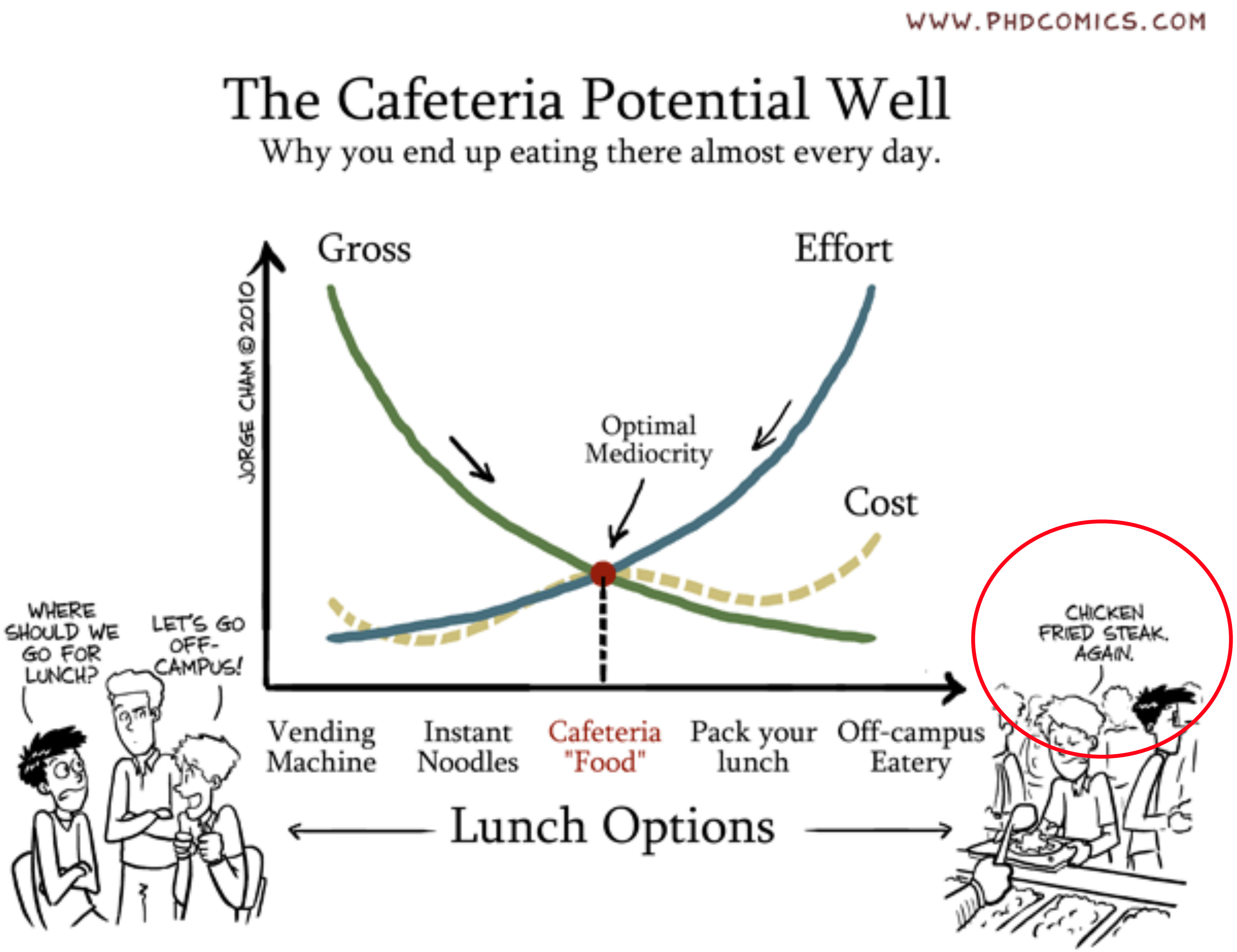
Ping-Chun Hsieh

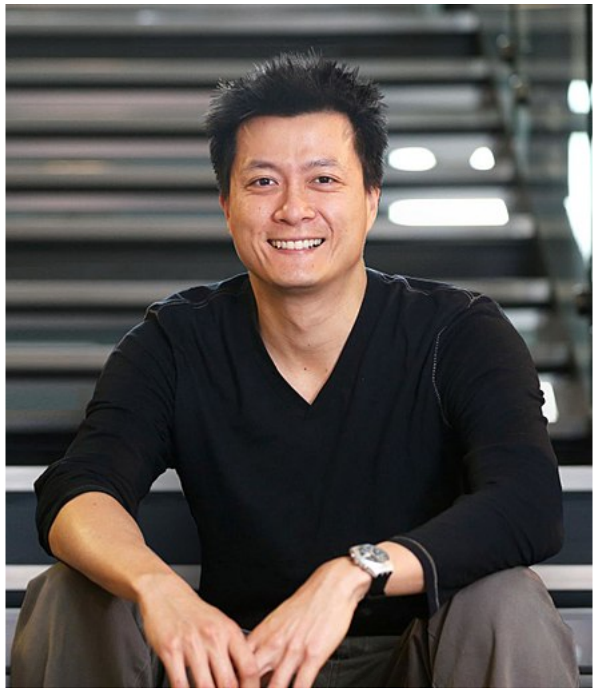
December 13, 2019

Announcements

- Class on 12/25 (Wed): 10:10am-**11:40**am
- Class on 12/13 (Fri): 3:30pm-**4:30**pm (10min extension)
- Class on 12/20 (Fri): 3:30pm-**4:30**pm (10min extension)

PHD Comics





Jorge Cham

http://phdcomics.com/comics/archive.php?comicid=1272

Quick Review

$Y_n \xrightarrow{P} Y$

Definition

$\lim_{n\to\infty} P(\{\omega : |Y_n(\omega) - Y(\omega)| > \varepsilon\}) = 0$ for every $\varepsilon > 0$

2 equivalent definitions

$Y_n \xrightarrow{a.s} Y$

Convergence in probability

implies

Almost-sure convergence

counter-example

$P(\{\omega : \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}) = 1$

WLLN

$X_1, X_2, \cdots, X_n, \cdots$ (mean $= \mu$)

$S_n = X_1 + X_2 + \cdots + X_n$

$\lim_{n\to\infty} P\left(\left\{\omega : \left|\frac{S_n(\omega)}{n} - \mu\right| > \varepsilon\right\}\right) = 0, \forall \varepsilon > 0$

SLLN

Convergence in Probability, But Not Almost Surely

$X_1(\omega) = \begin{cases} 1, & \text{if } X(\omega) \in [0,1] \\ 0, & \text{else} \end{cases}$

- Example: Let X be a continuous uniform r.v. on $(0,1)$
  - Consider a sequence of r.v.s $X_1, X_2, \cdots$ as follows:

$$X_1 = \mathbb{I}\{\boxed{X \in [0,1]}\}$$

$$X_2 = \mathbb{I}\{X \in [0,\frac{1}{2}]\} \qquad X_3 = \mathbb{I}\{X \in [\frac{1}{2},1]\}$$

$$X_4 = \mathbb{I}\{X \in [0,\frac{1}{3}]\} \qquad X_5 = \mathbb{I}\{X \in [\frac{1}{3},\frac{2}{3}]\} \qquad X_6 = \mathbb{I}\{X \in [\frac{2}{3},1]\}$$

$$\cdots \qquad \cdots \qquad \cdots$$

$P(X_1=1)=1$ $\quad$ $P(X_3=1)=0.5$

$P(X_2=1)=0.5$

$P(X_4=1)=\frac{1}{3}$ $\quad$ $P(X_5=1)=\frac{1}{3}$

$P(X_6=1)=\frac{1}{3}$

$P(X_{10000}=1) \approx 0$

- Question: Do we have $\lim_{n\to\infty} P(\{\omega : |X_n(\omega) - 0| > \varepsilon\}) = 0$?

$\varepsilon = 0.1$

Convergence in Probability, But Not Almost Surely (Cont.)

- Example: Let X be a continuous uniform r.v. on $(0,1)$
  - Consider a sequence of r.v.s $X_1, X_2, \cdots$ as follows:

$$X_1(\omega) = \mathbb{I}\{X(\omega) \in [0,1]\}$$

$$X_2(\omega) = \mathbb{I}\{X(\omega) \in [0,\frac{1}{2}]\} \quad X_3(\omega) = \mathbb{I}\{X(\omega) \in [\frac{1}{2},1]\}$$

$$X_4(\omega) = \mathbb{I}\{X(\omega) \in [0,\frac{1}{3}]\} \quad X_5(\omega) = \mathbb{I}\{X(\omega) \in [\frac{1}{3},\frac{2}{3}]\} \quad X_6(\omega) = \mathbb{I}\{X(\omega) \in [\frac{2}{3},1]\}$$

$$\cdots \qquad \cdots \qquad \cdots$$

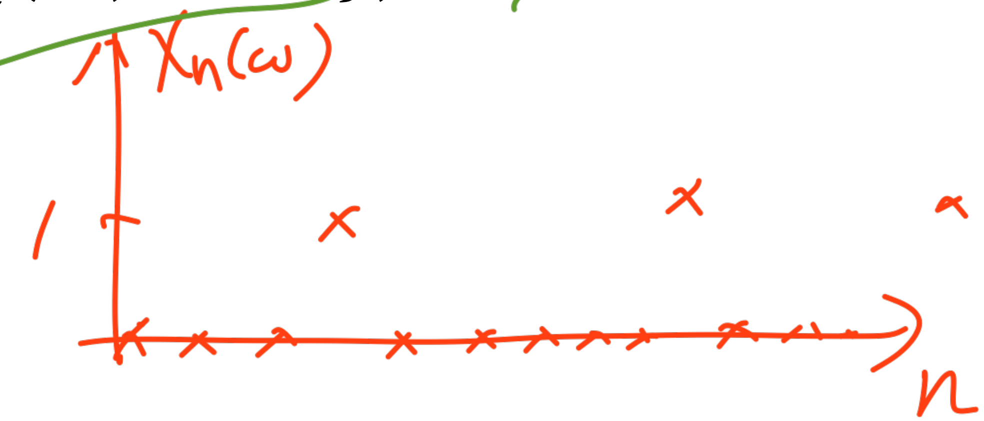
- Question: Do we have $P(\{\omega : \lim_{n\to\infty} X_n(\omega) = 0\}) = \cancel{1}\ 0$?

Fix an ω, which makes $X(\omega) = \cancel{0.18}\ 0.23$

$X_1(\omega) = 1$

$X_2(\omega) = 1$, $X_3(\omega) = 0$

$X_4(\omega) = 1$, $X_5(\omega) = 0$, $X_6(\omega) = 0$



This Lecture

1. Strong Law of Large Numbers (SLLN)

2. Central Limit Theorem (CLT)

- Reading material: Chapter 11.4-11.5

1. Strong Law of Large Numbers (SLLN)

WLLN vs SLLN

- **The Weak Law of Large Numbers (Khinchin's Law)**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ. Define $S_n = (X_1 + \cdots + X_n)$. Then, for every $\varepsilon > 0$, we have

$$\lim_{n\to\infty} P\Big(\{\omega : |\frac{S_n(\omega)}{n} - \mu| \geq \varepsilon\Big) = 0$$

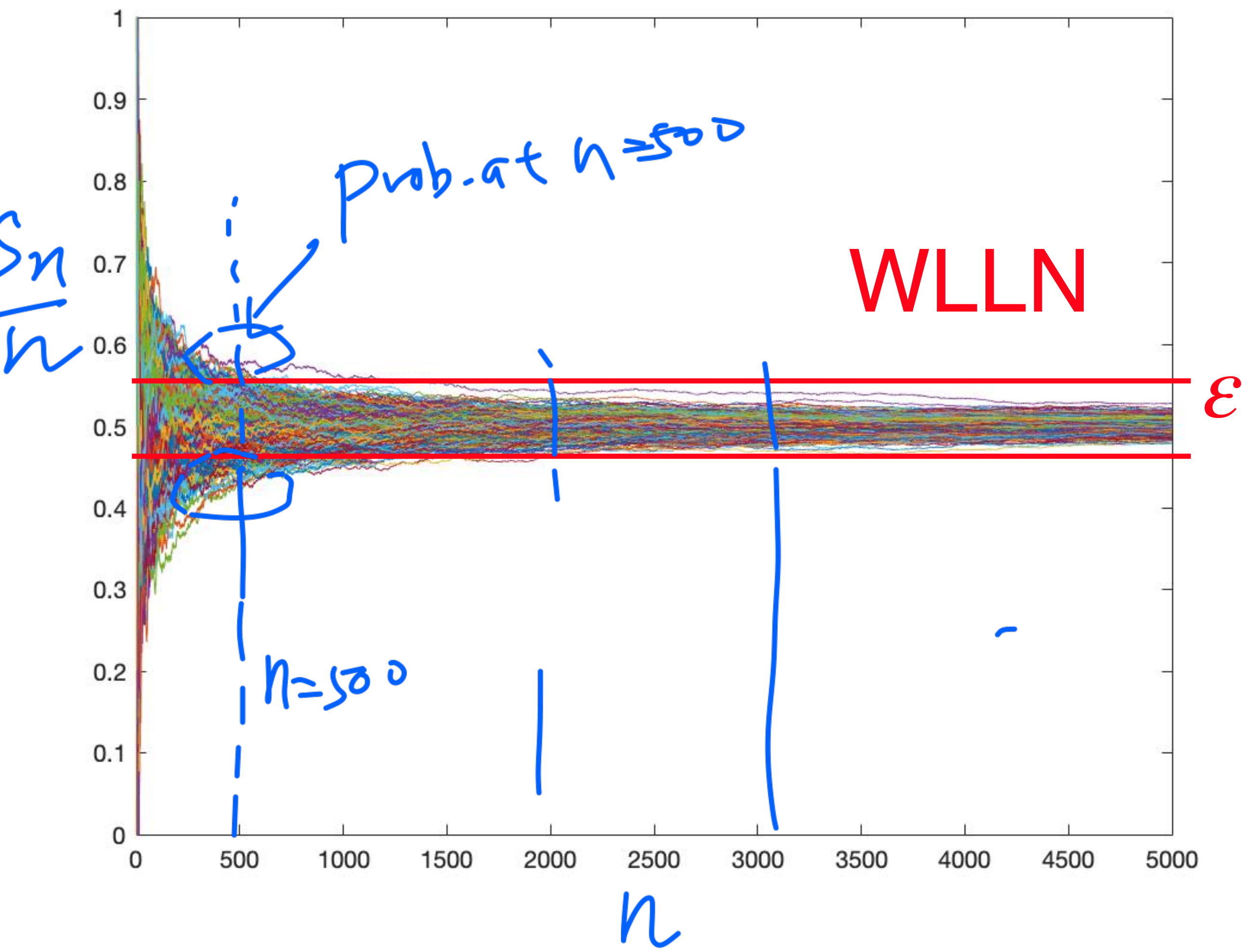
- **The Strong Law of Large Numbers**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ. Define $S_n = (X_1 + \cdots + X_n)$. Then, we have

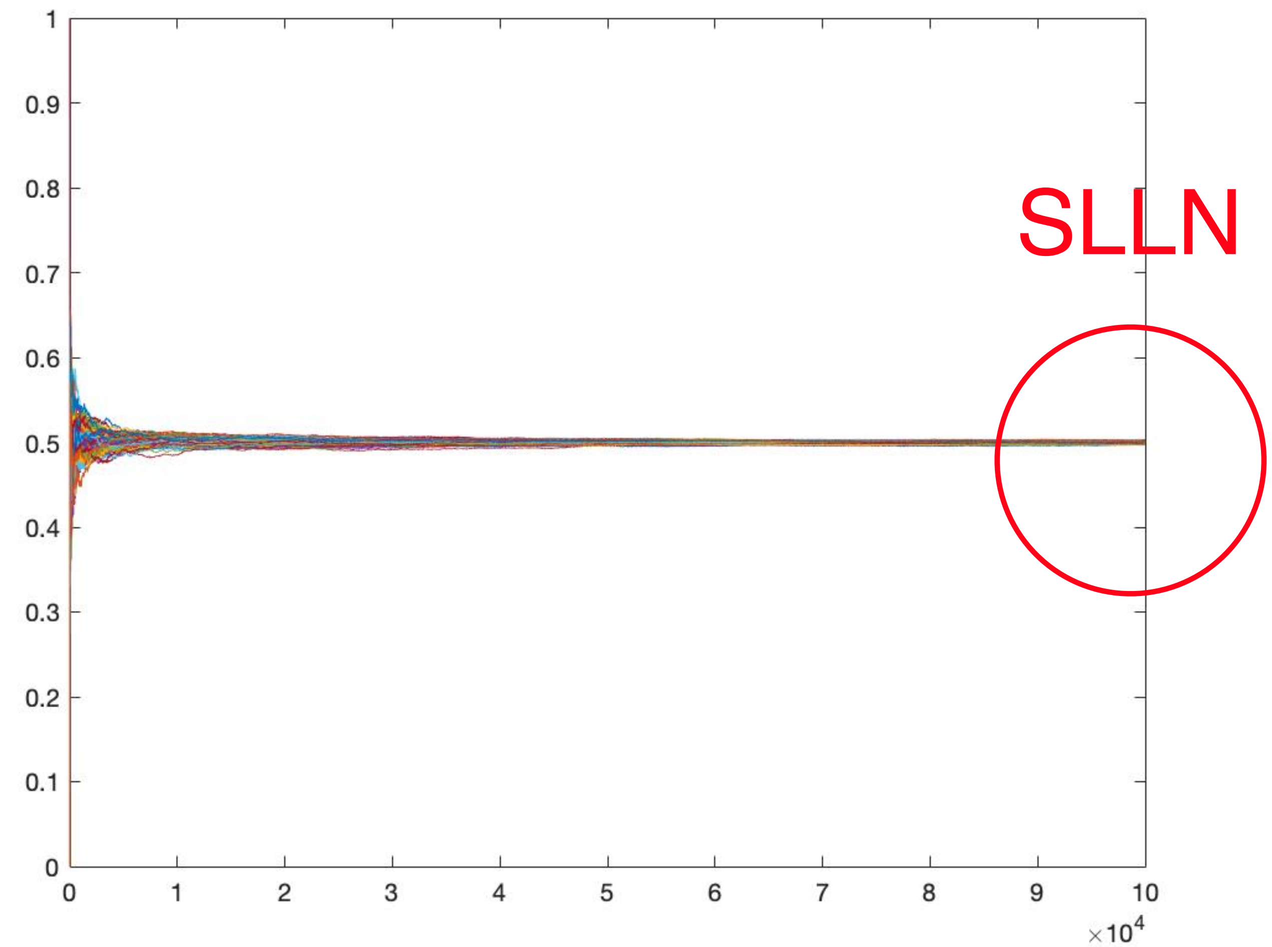
$$\frac{S_n}{n} \xrightarrow{a.s.} \mu \qquad P\Big(\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu\}\Big) = 1$$

(almost-sure convergence)

Visualization of WLLN and SLLN

- Example: $X_i \sim$ Bernoulli(0.5) and $S_n = X_1 + \cdots + X_n$



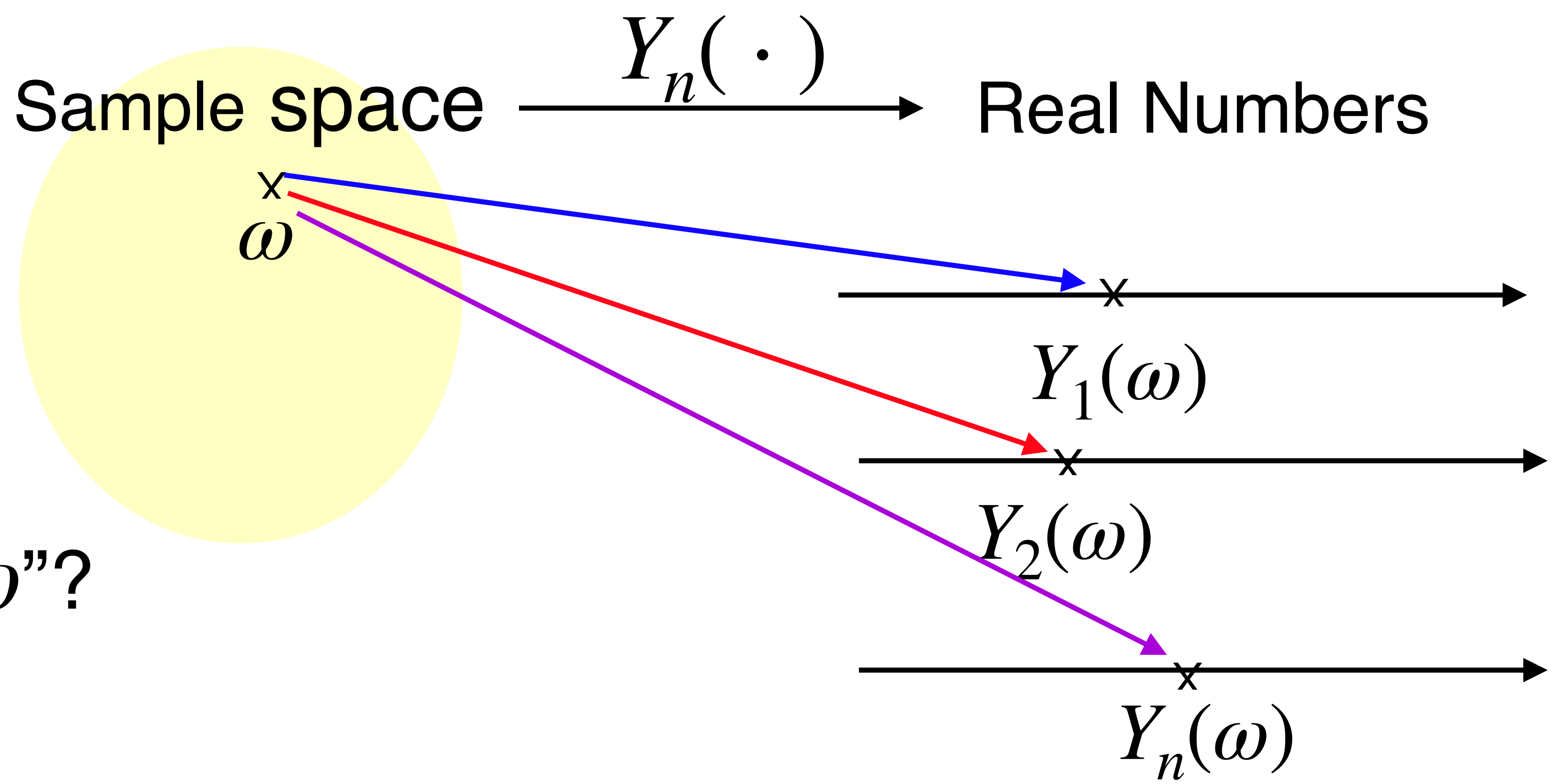


$$\lim_{n\to\infty} P\left(\{\omega : |\frac{S_n(\omega)}{n} - \mu| \geq \varepsilon\right) = 0$$

$$P\left(\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu\}\right) = 1$$

How to Interpret SLLN?

- Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables with mean μ
- Define $Y_n = (X_1 + X_2 \cdots, + X_n)/n$
- SLLN: $P\left(\left\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu\right\}\right) = 1$



- Question: What is an "ω"?

How to Prove SLLN (Under a Mild Condition)?

$E[X_i^4] < \infty$

(bounded 4-th moment)

1. Borel-Cantelli Lemma

2. A Bound for the 4-th Moment Condition

3. Markov's Inequality

1. Borel-Cantelli Lemma

- Recall: HW1, Problem 4

Problem 4 (Continuity of Probability Function and Probability Axioms) (8+8=16 points)

Consider an infinite sequence of coin tosses. The probability of having a head at the i-th toss is p_i, with $p_i \in [0,1]$ (Note: different tosses may not be independent and can potentially have different head probabilities). We use I to denote the event of having infinite number of heads.

(a) Suppose $\sum_{i=1}^{\infty} p_i$ is finite. Show that $P(I) = 0$. (Hint: Define $A_n := \{\text{the } n\text{-th toss is a head}\}$. Then, $B_k := \bigcup_{n=k}^{\infty} A_n$ is the event that there is at least one head after the k-th toss (including the k-th toss). The event I (i.e. we observe infinitely many heads) is equivalent to saying that B_k occurs, for every $k = 1, 2, 3, \cdots$. Therefore, $I = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$. Consider the continuity of probability function for $\bigcap_{k=1}^{\infty} \bigcup_{i=k}^{\infty} A_n$ and then apply the union bound)

- **Borel-Cantelli Lemma**: Let $\{A_n\}$ be any sequence of events. If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then we have

$$P\Big(A_n \text{ occurs infinitely often}\Big) = P\Big(\bigcap_{n\geq 1} \bigcup_{m\geq n} A_n\Big) = 0$$

Review: Proof of Borel-Cantelli Lemma

- **Borel-Cantelli Lemma**: Let $\{A_n\}$ be any sequence of events. If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then we have

$$P\Big(A_n \text{ occurs infinitely often}\Big) = P(\bigcap_{n\geq 1} \bigcup_{m \geq n} A_n) = 0$$

- Proof:

$$P\Big(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n\Big) = P\Big(\lim_{k\to\infty} \bigcap_{m=1}^{k}\bigcup_{n=m}^{\infty} A_n\Big)$$

decreasing

$$\bigcap_{m=1}^{k}\bigcup_{n=m}^{\infty} A_n = \bigcup_{n=k}^{\infty} A_n$$

$$= \lim_{k\to\infty} P\Big(\bigcap_{m=1}^{k}\bigcup_{n=m}^{\infty} A_n\Big)$$

$$= \lim_{k\to\infty} P\Big(\bigcup_{n=k}^{\infty} A_n\Big) \leq \lim_{k\to\infty} \sum_{n=k}^{\infty} P(A_n) = 0$$

2. A Bound For 4-th Moment of $\frac{S_n}{n}$

- **A Bound on 4-th Moment**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ and $E[X_1^4] < \infty$. Define $S_n = (X_1 + \cdots + X_n)$. Then, there exists a constant $K < \infty$ such that

$$\boxed{E[(S_n - n\mu)^4] \leq Kn^2}$$

- Question: How about $E[(\frac{S_n}{n} - \mu)^4] \leq ?$ $\frac{Kn^2}{n^4} = \frac{K}{n^2}$

$$\| $$

$$E\left[\frac{1}{n^4}(S_n - n\mu)^4\right]$$

Proof: A Bound For 4-th Moment

- Given: $S_n = (X_1 + \cdots + X_n)$ and $E[X_1^4] < \infty$
- Want: $E[(S_n - n\mu)^4] \leq Kn^2$
- Proof: For simplicity, let $Z_i = X_i - \mu$

Put Everything Together: Proof of SLLN

$$P\left(\left\{\left|\frac{S_n}{n}-\mu\right| \ge n^{-\gamma}\right\}\right) = P\left(\left\{\left|\frac{S_n}{n}-\mu\right|^4 \ge n^{-4\gamma}\right\}\right) \le \frac{E\left[\left(\frac{S_n}{n}-\mu\right)^4\right]}{n^{-4\gamma}}$$

by Markov's

$= A_n$

> 0

the 4-th moment bound

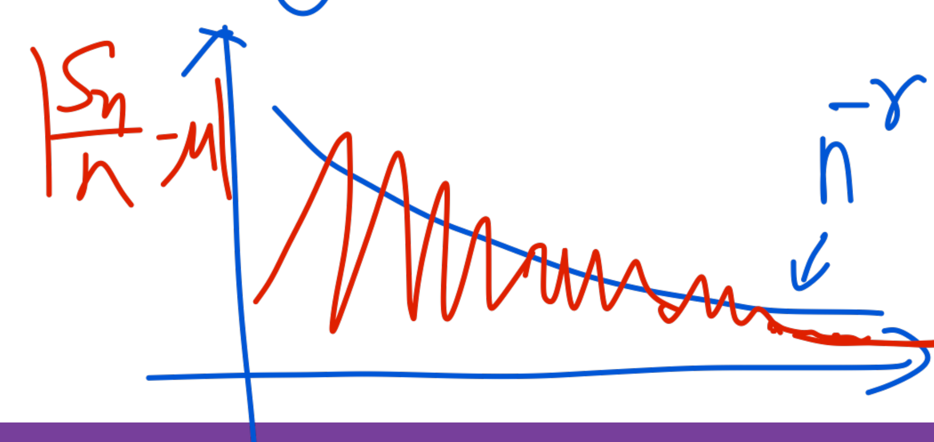
$$\le \frac{\frac{K}{n^2}}{n^{-4\gamma}} = K\cdot n^{-2+4\gamma}$$

Define events $A_n = \left\{ \left|\frac{S_n}{n}-\mu\right| \ge n^{-\gamma} \right\}$.

$$\sum_{n=1}^{\infty} P(A_n) \le \sum_{n=1}^{\infty} K\cdot n^{-2+4\gamma} < \infty, \text{ for any } \gamma < \frac{1}{4}$$

By Borel-Cantelli Lemma:

$P(A_n \text{ occurs infinitely often}) = 0$



2. Central Limit Theorem (CLT)

Beyond SLLN

- **The Strong Law of Large Numbers**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ. Define $S_n = (X_1 + \cdots + X_n)$. Then, we have
$$P\left(\left\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu\right\}\right) = 1$$

- Question: What does SLLN say about $S_n(\omega)$?

- Question: Do we have $S_n(\omega) = n\mu + o(n)$?

Review (Lecture 11): Binomial and Normal

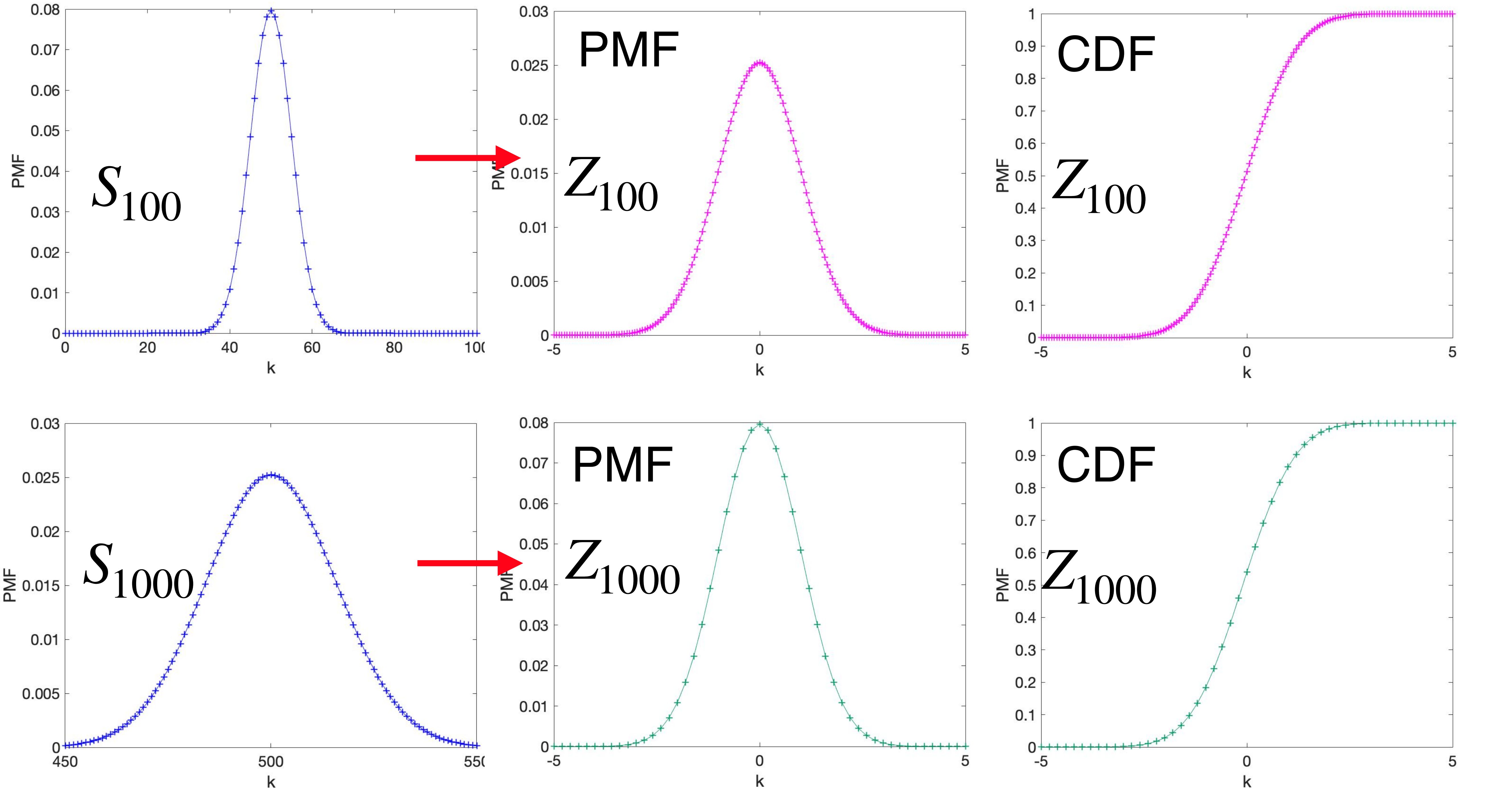
- Example: $X_1, X_2, \cdots$ are i.i.d. Bernoulli r.v.s with mean μ and variance $\sigma^2 = \mu(1-\mu)$
  - Define $S_n = X_1 + X_2 + \cdots + X_n$
  - Question: What type of r.v. is S_n? $E[S_n]$ = ? Var$[S_n]$ = ?

- Question: How to find the distribution of $\dfrac{S_n - n\mu}{\sigma\sqrt{n}}$?

Method 1: Plotting $Z_n = (S_n - n\mu)/(\sigma\sqrt{n})$

- Example: $\mu = 0.5$

$\approx$ CDF of a standard normal!



Method 2: MGF

- Idea: Suppose we find the MGF of $\dfrac{S_n - n\mu}{\sigma\sqrt{n}}$ for $n \to \infty$
  - Question: Can we find its distribution?

- **Levy Continuity Theorem**: Let $V_1, V_2 \cdots$ be a sequence of random variables with CDFs $F_1, F_2, \cdots$ and MGFs $M_{V_1}(t), M_{V_2}(t) \cdots$. Let V be a random variable with CDF F and MGF $M_V(t)$. If for every $t \in \mathbb{R}$, $\lim_{n \to \infty} M_{V_n}(t) = M_V(t)$, then the CDFs F_n converge to F.

Method 2: MGF (Cont.)

- Example: $X_1, X_2, \cdots$ are i.i.d. r.v.s with mean μ and variance σ^2
  - Define $S_n = X_1 + X_2 + \cdots + X_n$ and $Y_i = X_i - \mu$
  - Question: $E[Y_i]$ = ________? Var$[Y_i]$ = ________?
  - Question: What is the MGF of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ (in terms of MGF of Y_i)?

Method 2: MGF (Cont.)

- Question: When $n \to \infty$, what is the MGF of $\dfrac{S_n - n\mu}{\sigma\sqrt{n}}$?

Central Limit Theorem (Formally)

- **Central Limit Theorem (CLT)**: Let $X_1, \cdots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ and variance σ^2. Define $S_n = (X_1 + \cdots + X_n)$. Then, we have

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq z\right) = \Phi(z)$$

where $\Phi(z)$ is the CDF of standard normal

Next Lecture

- Parameter Estimation
  - Maximum Likelihood Estimation (MLE)
  - Maximum a Posteriori (MAP)

1-Minute Summary

1. Strong Law of Large Numbers (SLLN)

- SLLN: $P\left(\left\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu\right\}\right) = 1$
- Combo of 3 tools: Borel-Cantelli Lemma / Bound for 4th moment / Markov's inequality

2. Central Limit Theorem (CLT)

- CLT: $\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq z\right) = \Phi(z)$
- Proof by MGF