# Recap: Convergence in Probability

- A sequence of random variables, $X_n$, is said to **converge in probability** to a random variable $X_0$ is

$$\lim_{n \to \infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

# Recap: Convergence in Probability

- A sequence of random variables, $X_n$, is said to **converge in probability** to a random variable $X_0$ is

$$\lim_{n\to\infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

This is denoted as $X_n \stackrel{P}{\to} X_0$

# Recap: Convergence in Probability

- A sequence of random variables, $X_n$, is said to **converge in probability** to a random variable $X_0$ is

$$\lim_{n \to \infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

  This is denoted as $X_n \xrightarrow{P} X_0$

- By the definition of limit, the above means

$$\forall \delta > 0, \ \exists N < \infty, \ s.t. \ P\left[|X_n - X_0| > \epsilon\right] < \delta, \ \forall n > N$$

# Recap: Convergence in Probability

- A sequence of random variables, $X_n$, is said to **converge in probability** to a random variable $X_0$ is

$$\lim_{n\to\infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

  This is denoted as $X_n \stackrel{P}{\to} X_0$
- By the definition of limit, the above means

$$\forall \delta > 0, \ \exists N < \infty, \ s.t. \ P\left[|X_n - X_0| > \epsilon\right] < \delta, \ \forall n > N$$

- We only need marginal distributions of individual $X_n$ to decide whether a sequence converges to a constant in probability

# Recap: Weak Law of large numbers

- $X_i$ are iid, $EX_i = \mu$, $\text{Var}(X_i) = \sigma^2$, $S_n = \sum_{i=1}^{n} X_i$

# Recap: Weak Law of large numbers

- $X_i$ are iid, $EX_i = \mu$, $\text{Var}(X_i) = \sigma^2$, $S_n = \sum_{i=1}^{n} X_i$

$$E\left[\frac{S_n}{n}\right] = \mu; \quad \text{and} \quad \text{Var}\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}$$

# Recap: Weak Law of large numbers

- $X_i$ are iid, $EX_i = \mu$, $\text{Var}(X_i) = \sigma^2$, $S_n = \sum_{i=1}^n X_i$

$$E\left[\frac{S_n}{n}\right] = \mu; \quad \text{and} \quad \text{Var}\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}$$

Weak law of large numbers states

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

# Recap: almost sure convergence

- A sequence of random variables, $X_n$, is said to converge **almost surely** or **with probability one** to $X$ if

$$P\left(\{\omega \ : \ X_n(\omega) \to X(\omega)\}\right) = 1$$

# Recap: almost sure convergence

- A sequence of random variables, $X_n$, is said to converge **almost surely** or **with probability one** to $X$ if

$$P\left(\{\omega \ : \ X_n(\omega) \rightarrow X(\omega)\}\right) = 1$$

or equivalently

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

# Recap: almost sure convergence

- A sequence of random variables, $X_n$, is said to converge **almost surely** or **with probability one** to $X$ if

$$P\left(\{\omega \ : \ X_n(\omega) \to X(\omega)\}\right) = 1$$

or equivalently

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

- Denoted as $X_n \stackrel{a.s.}{\to} X$ or $X_n \stackrel{w.p.1}{\to} X$ or $X_n \to X_0 \ \ (w.p.1)$

# Recap: almost sure convergence

- A sequence of random variables, $X_n$, is said to converge **almost surely** or **with probability one** to $X$ if

$$P\left(\{\omega\ :\ X_n(\omega) \rightarrow X(\omega)\}\right) = 1$$

  or equivalently

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

- Denoted as $X_n \stackrel{a.s.}{\rightarrow} X$ or $X_n \stackrel{w.p.1}{\rightarrow} X$ or $X_n \rightarrow X_0 \ \ (w.p.1)$
- We can also write it as

$$P[X_n \rightarrow X] = 1$$

# Recap

- The sequence $X_n$ converges to $X$ almost surely iff
$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \; [|X_{N+k} - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

# Recap

- The sequence $X_n$ converges to $X$ almost surely iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \; [|X_{N+k} - X| \geq \epsilon]\right) = 0, \;\; \forall \epsilon > 0$$

Same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \; [|X_k - X| \geq \epsilon]\right) = 0, \;\; \forall \epsilon > 0$$

# Recap

- The sequence $X_n$ converges to $X$ almost surely iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \ [|X_{N+k} - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

  Same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- Equivalently

$$\lim_{n \to \infty} \ P\left(\cup_{k=n}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

# Recap

- The sequence $X_n$ converges to $X$ almost surely iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty}\ [|X_{N+k} - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

  Same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty}\ [|X_k - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

- Equivalently

$$\lim_{n\to\infty} P\left(\cup_{k=n}^{\infty}\ [|X_k - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

- $X_n \xrightarrow{P} X$ iff

$$\lim_{n\to\infty} P\left[|X_n - X| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

# Recap

- The sequence $X_n$ converges to $X$ almost surely iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \ [|X_{N+k} - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

  Same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

- Equivalently

$$\lim_{n\to\infty} \ P\left(\cup_{k=n}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

- $X_n \stackrel{P}{\to} X$ iff

$$\lim_{n\to\infty} P\left[|X_n - X| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

$$X_n \stackrel{a.s.}{\to} X \quad \Rightarrow \quad X_n \stackrel{P}{\to} X$$

# Recap: lim sup and lim inf

- Let $A_1, A_2, \cdots$ be a sequence of events.

# Recap: lim sup and lim inf

- Let $A_1, A_2, \cdots$ be a sequence of events.
- We define

$$\lim\ \sup A_n \ \triangleq\ \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$$

# Recap: lim sup and lim inf

- Let $A_1, A_2, \cdots$ be a sequence of events.
- We define

$$\begin{aligned} \lim\ \sup A_n &\triangleq \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \\ \lim\ \inf A_n &\triangleq \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k \end{aligned}$$

# Recap: lim sup and lim inf

- Let $A_1, A_2, \cdots$ be a sequence of events.
- We define

$$\lim\ \sup A_n \triangleq \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$$
$$\lim\ \inf A_n \triangleq \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k$$

- If $\lim\ \sup A_n = \lim\ \inf A_n$ then that is $\lim A_n$. Otherwise the sequence does not have a limit

# Recap: lim sup and lim inf

- Let $A_1, A_2, \cdots$ be a sequence of events.
- We define

$$\begin{aligned} \lim\ \sup A_n &\triangleq \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \\ \lim\ \inf A_n &\triangleq \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k \end{aligned}$$

- If $\lim\ \sup A_n = \lim\ \inf A_n$ then that is $\lim A_n$. Otherwise the sequence does not have a limit
- $\lim\ \sup A_n$ and $\lim\ \inf A_n$ are events

# Recap: lim sup and lim inf

- Let $A_1, A_2, \cdots$ be a sequence of events.
- We define

$$\lim\ \sup A_n \triangleq \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$$
$$\lim\ \inf A_n \triangleq \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k$$

- If $\lim\ \sup A_n = \lim\ \inf A_n$ then that is $\lim A_n$. Otherwise the sequence does not have a limit
- $\lim\ \sup A_n$ and $\lim\ \inf A_n$ are events
- $\lim\ \inf A_n \subset \lim\ \sup A_n$

# Recap

- $X_n \stackrel{a.s.}{\rightarrow} X$ iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \; [|X_k - X| \geq \epsilon]\right) = 0, \;\; \forall \epsilon > 0$$

# Recap

- $X_n \stackrel{a.s.}{\rightarrow} X$ iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$.

# Recap

- $X_n \stackrel{a.s.}{\to} X$ iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \; [|X_k - X| \geq \epsilon]\right) = 0, \;\; \forall \epsilon > 0$$

- Let $A_k^{\epsilon} = [|X_k - X| \geq \epsilon]$.
- Hence, $X_n \stackrel{a.s.}{\to} X$ iff

$$P\left(\lim\ \sup\ A_n^{\epsilon}\right) = 0, \;\; \forall \epsilon > 0$$

# Recall: Borel-Cantelli Lemma

# Recall: Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given sequence of events, $A_1, A_2, \cdots$

# Recall: Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given sequence of events, $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$

# Recall: Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given sequence of events, $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

# Proof of Borel-Cantelli Lemma

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i), \ \forall n$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i), \ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \leq \sum_{i=n}^{N} P(A_i), n \leq N$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \le \sum_{i=n}^{\infty} P(A_i), \ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \le \sum_{i=n}^{N} P(A_i), n \le N$
- For any $n$, let $B_N = \cup_{i=n}^{N} A_i$. Then $B_N \subset B_{N+1}$.

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i),\ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \leq \sum_{i=n}^{N} P(A_i), n \leq N$
- For any $n$, let $B_N = \cup_{i=n}^{N} A_i$. Then $B_N \subset B_{N+1}$.
- $\lim_{N\to\infty} B_N = \cup_{k=n}^{\infty} A_k$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i), \ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \leq \sum_{i=n}^{N} P(A_i), n \leq N$
- For any $n$, let $B_N = \cup_{i=n}^{N} A_i$. Then $B_N \subset B_{N+1}$.
- $\lim_{N\to\infty} B_N = \cup_{k=n}^{\infty} A_k$

$$P(\cup_{i=n}^{\infty} A_i) = P(\lim_{N\to\infty} \cup_{i=n}^{N} A_i)$$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i), \ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \leq \sum_{i=n}^{N} P(A_i), n \leq N$
- For any $n$, let $B_N = \cup_{i=n}^{N} A_i$. Then $B_N \subset B_{N+1}$.
- $\lim_{N\to\infty} B_N = \cup_{k=n}^{\infty} A_k$

$$P(\cup_{i=n}^{\infty} A_i) = P(\lim_{N\to\infty} \cup_{i=n}^{N} A_i) \quad = \quad \lim_{N\to\infty} P(\cup_{i=n}^{N} A_i)$$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

  **Proof:**
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \le \sum_{i=n}^{\infty} P(A_i), \ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \le \sum_{i=n}^{N} P(A_i), n \le N$
- For any $n$, let $B_N = \cup_{i=n}^{N} A_i$. Then $B_N \subset B_{N+1}$.
- $\lim_{N\to\infty} B_N = \cup_{k=n}^{\infty} A_k$

$$
\begin{aligned}
P(\cup_{i=n}^{\infty} A_i) = P(\lim_{N\to\infty} \cup_{i=n}^{N} A_i) &= \lim_{N\to\infty} P(\cup_{i=n}^{N} A_i) \\
&\le \lim_{N\to\infty} \sum_{i=n}^{N} P(A_i)
\end{aligned}
$$

# Proof of Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and $A_i$ are independent, $P(\lim\ \sup A_n) = 1$

**Proof:**

- We will first show: $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i), \ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \leq \sum_{i=n}^{N} P(A_i), n \leq N$
- For any $n$, let $B_N = \cup_{i=n}^{N} A_i$. Then $B_N \subset B_{N+1}$.
- $\lim_{N\to\infty} B_N = \cup_{k=n}^{\infty} A_k$

$$
\begin{aligned}
P(\cup_{i=n}^{\infty} A_i) = P(\lim_{N\to\infty} \cup_{i=n}^{N} A_i) &= \lim_{N\to\infty} P(\cup_{i=n}^{N} A_i) \\
&\leq \lim_{N\to\infty} \sum_{i=n}^{N} P(A_i) = \sum_{i=n}^{\infty} P(A_i)
\end{aligned}
$$

- By definition,
  $\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

- By definition, $\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n \to \infty} \sum_{k=n}^{\infty} P(A_k) = 0$
- Let $\sum_{k=1}^{\infty} P(A_k) = C < \infty$

- By definition, $\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$
- Let $\sum_{k=1}^{\infty} P(A_k) = C < \infty$
- It means given any $\epsilon > 0$, $\exists n$

$$\left| \sum_{k=1}^{n} P(A_k) - C \right| < \epsilon$$

- By definition, $\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$
- Let $\sum_{k=1}^{\infty} P(A_k) = C < \infty$
- It means given any $\epsilon > 0$, $\exists n$

$$\left| \sum_{k=1}^{n} P(A_k) - C \right| < \epsilon \;\; \Rightarrow \;\; \left| \sum_{k=n}^{\infty} P(A_k) \right| < \epsilon$$

- By definition, $\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$
- Let $\sum_{k=1}^{\infty} P(A_k) = C < \infty$
- It means given any $\epsilon > 0$, $\exists n$

$$\left| \sum_{k=1}^{n} P(A_k) - C \right| < \epsilon \quad \Rightarrow \quad \left| \sum_{k=n}^{\infty} P(A_k) \right| < \epsilon$$

- This implies

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$$

- By definition, $\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$
- Let $\sum_{k=1}^{\infty} P(A_k) = C < \infty$
- It means given any $\epsilon > 0$, $\exists n$

$$\left| \sum_{k=1}^{n} P(A_k) - C \right| < \epsilon \quad \Rightarrow \quad \left| \sum_{k=n}^{\infty} P(A_k) \right| < \epsilon$$

- This implies

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$$

- Similarly we have $\sum_{k=1}^{\infty} P(A_k) = \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = \infty$

- By definition,
$\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

- By definition,
$\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

$$0 \leq P\left(\lim\ \sup A_n\right) \;=\; P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right)$$

- By definition,
$\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

$$
\begin{aligned}
0 \le P(\lim\ \sup A_n) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right)
\end{aligned}
$$

- By definition,
$\sum_{k=1}^{\infty} P(A_k) < \infty$, $\Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

$$\begin{aligned} 0 \le P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\ &= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\ &= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \end{aligned}$$

- By definition,
$\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

$$\begin{aligned} 0 \leq P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\ &= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\ &= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\ &\leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) \end{aligned}$$

- By definition,
$\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

$$\begin{aligned}
0 \le P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&\le \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) \\
&= 0, \quad \text{if} \quad \sum_{k=1}^{\infty} P(A_k) < \infty
\end{aligned}$$

- By definition, $\sum_{k=1}^{\infty} P(A_k) < \infty, \Rightarrow \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

$$
\begin{aligned}
0 \le P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&\le \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) \\
&= 0, \quad \text{if} \quad \sum_{k=1}^{\infty} P(A_k) < \infty
\end{aligned}
$$

- This completes proof of first part of Borel-Cantelli lemma

- For the second part of the lemma:

- For the second part of the lemma:

$$P\left(\lim\ \sup A_n\right) = P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right)$$

- For the second part of the lemma:

$$
\begin{aligned}
P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right)
\end{aligned}
$$

- For the second part of the lemma:

$$\begin{aligned} P(\lim\ \sup A_n) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\ &= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\ &= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \end{aligned}$$

- For the second part of the lemma:

$$
\begin{aligned}
P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} \left(1 - P\left(\cap_{k=n}^{\infty} A_k^c\right)\ \right)
\end{aligned}
$$

- For the second part of the lemma:

$$
\begin{aligned}
P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} \left(1 - P\left(\cap_{k=n}^{\infty} A_k^c\right)\right) \\
&= \lim_{n\to\infty} \left(1 - \prod_{k=n}^{\infty} (1 - P(A_k))\right)
\end{aligned}
$$

- For the second part of the lemma:

$$
\begin{aligned}
P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} \left(1 - P\left(\cap_{k=n}^{\infty} A_k^c\right)\ \right) \\
&= \lim_{n\to\infty} \left(1 - \prod_{k=n}^{\infty} (1 - P(A_k))\ \right) \\
& \quad \text{because } A_k \text{ are independent}
\end{aligned}
$$

- For the second part of the lemma:

$$\begin{aligned}
P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} \left(1 - P\left(\cap_{k=n}^{\infty} A_k^c\right)\ \right) \\
&= \lim_{n\to\infty} \left(1 - \prod_{k=n}^{\infty} (1 - P(A_k))\ \right) \\
&\quad \text{because } A_k \text{ are independent} \\
&= 1\ -\ \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k))
\end{aligned}$$

- We can compute that limit as follows

- We can compute that limit as follows

$$\lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) \;\; \leq \;\; \lim_{n\to\infty} \prod_{k=n}^{\infty} e^{-P(A_k)}, \quad \text{since} \;\; 1 - x \leq e^{-x}$$

- We can compute that limit as follows

$$
\begin{aligned}
\lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) &\leq \lim_{n\to\infty} \prod_{k=n}^{\infty} e^{-P(A_k)}, \quad \text{since } \ 1 - x \leq e^{-x} \\
&= \lim_{n\to\infty} \ e^{-\sum_{k=n}^{\infty} P(A_k)}
\end{aligned}
$$

- We can compute that limit as follows

$$\begin{aligned}
\lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) &\le \lim_{n\to\infty} \prod_{k=n}^{\infty} e^{-P(A_k)}, \quad \text{since } \ 1 - x \le e^{-x} \\
&= \lim_{n\to\infty} \ e^{-\sum_{k=n}^{\infty} P(A_k)} \\
&= 0
\end{aligned}$$

- We can compute that limit as follows

$$
\begin{aligned}
\lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) &\leq \lim_{n\to\infty} \prod_{k=n}^{\infty} e^{-P(A_k)}, \quad \text{since } \ 1 - x \leq e^{-x} \\
&= \lim_{n\to\infty} \ e^{-\sum_{k=n}^{\infty} P(A_k)} \\
&= 0
\end{aligned}
$$

because
$\sum_{k=1}^{\infty} P(A_k) = \infty \quad \Rightarrow \quad \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = \infty$

- We can compute that limit as follows

$$
\begin{aligned}
\lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) &\leq \lim_{n\to\infty} \prod_{k=n}^{\infty} e^{-P(A_k)}, \quad \text{since } \ 1 - x \leq e^{-x} \\
&= \lim_{n\to\infty} \ e^{-\sum_{k=n}^{\infty} P(A_k)} \\
&= 0
\end{aligned}
$$

because
$\sum_{k=1}^{\infty} P(A_k) = \infty \quad \Rightarrow \quad \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = \infty$

- This finally gives us

$$
P\left(\lim\ \sup A_n\right) = 1 \ - \ \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) = 1
$$

- We can compute that limit as follows

$$\begin{aligned}
\lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) &\leq \lim_{n\to\infty} \prod_{k=n}^{\infty} e^{-P(A_k)}, \quad \text{since} \;\; 1 - x \leq e^{-x} \\
&= \lim_{n\to\infty} \; e^{-\sum_{k=n}^{\infty} P(A_k)} \\
&= 0
\end{aligned}$$

because
$\sum_{k=1}^{\infty} P(A_k) = \infty \quad \Rightarrow \quad \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = \infty$

- This finally gives us

$$P\left(\lim\ \sup A_n\right) = 1 \ - \ \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)) = 1$$

- That completes the proof.

- Given a sequence $X_n$ we want to know whether it converges to $X$

- Given a sequence $X_n$ we want to know whether it converges to $X$
- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$

- Given a sequence $X_n$ we want to know whether it converges to $X$
- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$
- $X_n \xrightarrow{P} X$ if

$$\lim_{k\to\infty} P[|X_k - X| \geq \epsilon] = 0$$

- Given a sequence $X_n$ we want to know whether it converges to $X$
- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$
- $X_n \xrightarrow{P} X$ if

$$\lim_{k\to\infty} P[|X_k - X| \geq \epsilon] = 0 \quad \text{same as} \quad \lim_{k\to\infty} P(A_k^\epsilon) = 0, \;\; \forall \epsilon > 0$$

- Given a sequence $X_n$ we want to know whether it converges to $X$
- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$
- $X_n \xrightarrow{P} X$ if

$$\lim_{k\to\infty} P[|X_k - X| \geq \epsilon] = 0 \quad \text{same as} \quad \lim_{k\to\infty} P(A_k^\epsilon) = 0, \;\; \forall \epsilon > 0$$

- By Borel-Cantelli lemma, if $\forall \epsilon > 0$,

$$\sum_{k=1}^{\infty} P(A_k^\epsilon) < \infty$$

- Given a sequence $X_n$ we want to know whether it converges to $X$
- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$
- $X_n \xrightarrow{P} X$ if

$$\lim_{k\to\infty} P[|X_k - X| \geq \epsilon] = 0 \quad \text{same as} \quad \lim_{k\to\infty} P(A_k^\epsilon) = 0, \;\; \forall \epsilon > 0$$

- By Borel-Cantelli lemma, if $\forall \epsilon > 0$,

$$\sum_{k=1}^{\infty} P(A_k^\epsilon) < \infty \;\; \Rightarrow \;\; P(\lim\ \sup A_k^\epsilon) = 0 \;\; \Rightarrow \;\; X_k \xrightarrow{a.s.} X$$

- Given a sequence $X_n$ we want to know whether it converges to $X$
- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$
- $X_n \xrightarrow{P} X$ if

$$\lim_{k\to\infty} P[|X_k - X| \geq \epsilon] = 0 \quad \text{same as} \quad \lim_{k\to\infty} P(A_k^\epsilon) = 0, \;\; \forall \epsilon > 0$$

- By Borel-Cantelli lemma, if $\forall \epsilon > 0$,

$$\sum_{k=1}^{\infty} P(A_k^\epsilon) < \infty \;\; \Rightarrow \;\; P(\lim\ \sup A_k^\epsilon) = 0 \;\; \Rightarrow \;\; X_k \xrightarrow{a.s.} X$$

If $X_k$ are ind

$$\sum_{k=1}^{\infty} P(A_k^\epsilon) = \infty \;\; \Rightarrow \;\; P(\lim\ \sup A_k) = 1 \;\; \Rightarrow \;\; X_k \overset{a.s.}{\nrightarrow} X$$

# Recall Example

- Consider a sequence $X_n$ with

$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$

# Recall Example

- Consider a sequence $X_n$ with

$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$

- Since $\frac{1}{n} \to 0$ as $n \to \infty$, we conclude $X_n \xrightarrow{P} 0$.

# Recall Example

- Consider a sequence $X_n$ with
$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$
- Since $\frac{1}{n} \to 0$ as $n \to \infty$, we conclude $X_n \stackrel{P}{\to} 0$.
- But since, $\sum_n \frac{1}{n} = \infty$, Borel-Cantelli is not useful

# Recall Example

- Consider a sequence $X_n$ with
$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$
- Since $\frac{1}{n} \to 0$ as $n \to \infty$, we conclude $X_n \xrightarrow{P} 0$.
- But since, $\sum_n \frac{1}{n} = \infty$, Borel-Cantelli is not useful
- We saw one example of such $X_n$ with $X_k \xrightarrow{a.s.} 0$

# Recall Example

- Consider a sequence $X_n$ with

$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$

- Since $\frac{1}{n} \to 0$ as $n \to \infty$, we conclude $X_n \xrightarrow{P} 0$.
- But since, $\sum_n \frac{1}{n} = \infty$, Borel-Cantelli is not useful
- We saw one example of such $X_n$ with $X_k \xrightarrow{a.s.} 0$
- But, if $X_n$ are independent, then by Borel-Cantelli lemma, they do not converge

# Recall Example

- Consider a sequence $X_n$ with
$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$
- Since $\frac{1}{n} \rightarrow 0$ as $n \rightarrow \infty$, we conclude $X_n \stackrel{P}{\rightarrow} 0$.
- But since, $\sum_n \frac{1}{n} = \infty$, Borel-Cantelli is not useful
- We saw one example of such $X_n$ with $X_k \stackrel{a.s.}{\rightarrow} 0$
- But, if $X_n$ are independent, then by Borel-Cantelli lemma, they do not converge
- Almost sure convergence is stronger than convergence in probability

# Recall Example

- Consider a sequence $X_n$ with
$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$
- Since $\frac{1}{n} \to 0$ as $n \to \infty$, we conclude $X_n \xrightarrow{P} 0$.
- But since, $\sum_n \frac{1}{n} = \infty$, Borel-Cantelli is not useful
- We saw one example of such $X_n$ with $X_k \xrightarrow{a.s.} 0$
- But, if $X_n$ are independent, then by Borel-Cantelli lemma, they do not converge
- Almost sure convergence is stronger than convergence in probability
- Convergence (to a constant) in probability depends only on distribution of individual $X_n$.

# Recall Example

- Consider a sequence $X_n$ with
$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$
- Since $\frac{1}{n} \to 0$ as $n \to \infty$, we conclude $X_n \xrightarrow{P} 0$.
- But since, $\sum_n \frac{1}{n} = \infty$, Borel-Cantelli is not useful
- We saw one example of such $X_n$ with $X_k \xrightarrow{a.s.} 0$
- But, if $X_n$ are independent, then by Borel-Cantelli lemma, they do not converge
- Almost sure convergence is stronger than convergence in probability
- Convergence (to a constant) in probability depends only on distribution of individual $X_n$.
- Convergence almost surely depends on the joint distribution

# Strong Law of Large Numbers

- Let $X_n$ be iid, $EX_n = \mu$, $\mathsf{Var}(X_n) = \sigma^2$, $S_n = \sum_{i=1}^n X_i$

# Strong Law of Large Numbers

- Let $X_n$ be iid, $EX_n = \mu$, Var$(X_n) = \sigma^2$, $S_n = \sum_{i=1}^{n} X_i$
- We saw weak law of large numbers:

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

# Strong Law of Large Numbers

- Let $X_n$ be iid, $EX_n = \mu$, $\text{Var}(X_n) = \sigma^2$, $S_n = \sum_{i=1}^n X_i$
- We saw weak law of large numbers:

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

- Strong law of large numbers says:

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu$$

# Strong Law of Large Numbers

- Let $X_n$ be iid, $EX_n = \mu$, $\mathrm{Var}(X_n) = \sigma^2$, $S_n = \sum_{i=1}^{n} X_i$
- We saw weak law of large numbers:

$$\frac{S_n}{n} \stackrel{P}{\rightarrow} \mu$$

- Strong law of large numbers says:

$$\frac{S_n}{n} \stackrel{a.s.}{\rightarrow} \mu$$

- This is true for **all** random variables

# Strong Law of Large Numbers

- Let $X_n$ be iid, $EX_n = \mu$, $\mathrm{Var}(X_n) = \sigma^2$, $S_n = \sum_{i=1}^{n} X_i$
- We saw weak law of large numbers:

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

- Strong law of large numbers says:

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu$$

- This is true for **all** random variables
- We prove only a restricted version

- Let $A_n^\epsilon = \left[ \left| \frac{S_n}{n} - \mu \right| > \epsilon \right]$

- Let $A_n^\epsilon = \left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right]$
- As we saw, by Chebyshev inequality

$$P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] \leq \frac{\sigma^2}{n\epsilon^2}$$

- Let $A_n^\epsilon = \left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right]$
- As we saw, by Chebyshev inequality

$$P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] \leq \frac{\sigma^2}{n\epsilon^2}$$

- This shows $P(A_n^\epsilon) \to 0$ and thus we get weak law

- Let $A_n^\epsilon = \left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right]$
- As we saw, by Chebyshev inequality

$$P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] \leq \frac{\sigma^2}{n\epsilon^2}$$

- This shows $P(A_n^\epsilon) \to 0$ and thus we get weak law
- To prove strong law using Borel-Cantelli lemma, we need $\sum P(A_n^\epsilon) < \infty$

- Let $A_n^\epsilon = \left[ \left| \frac{S_n}{n} - \mu \right| > \epsilon \right]$
- As we saw, by Chebyshev inequality

$$P\left[ \left| \frac{S_n}{n} - \mu \right| > \epsilon \right] \leq \frac{\sigma^2}{n\epsilon^2}$$

- This shows $P(A_n^\epsilon) \to 0$ and thus we get weak law
- To prove strong law using Borel-Cantelli lemma, we need $\sum P(A_n^\epsilon) < \infty$
- Since $\sum_n \frac{\sigma^2}{n\epsilon^2} = \infty$, the Chebyshev bound is not useful

- Let $A_n^\epsilon = \left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right]$
- As we saw, by Chebyshev inequality

$$P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] \leq \frac{\sigma^2}{n\epsilon^2}$$

- This shows $P(A_n^\epsilon) \to 0$ and thus we get weak law
- To prove strong law using Borel-Cantelli lemma, we need $\sum P(A_n^\epsilon) < \infty$
- Since $\sum_n \frac{\sigma^2}{n\epsilon^2} = \infty$, the Chebyshev bound is not useful
- We need a bound: $P[|\frac{S_n}{n} - \mu|] \leq c_n$ such that $\sum_n c_n < \infty$.

- Let us assume $X_i$ have finite fourth moment

- Let us assume $X_i$ have finite fourth moment

$$\left(\sum_{i=1}^{n}(X_i - \mu)\right)^4 =$$

- Let us assume $X_i$ have finite fourth moment

$$\left(\sum_{i=1}^{n}(X_i-\mu)\right)^4 = \sum_{i=1}^{n}(X_i-\mu)^4+\sum_{i}\sum_{j>i}\ \frac{4!}{2!2!}\ (X_i-\mu)^2(X_j-\mu)^2+T$$

- Let us assume $X_i$ have finite fourth moment

$$\left(\sum_{i=1}^{n}(X_i - \mu)\right)^4 = \sum_{i=1}^{n}(X_i - \mu)^4 + \sum_{i}\sum_{j>i} \frac{4!}{2!2!}(X_i - \mu)^2 (X_j - \mu)^2 + T$$

Where $T$ represent a number of terms such that every term in it contains a factor like $(X_i - \mu)$

- Let us assume $X_i$ have finite fourth moment

$$\left(\sum_{i=1}^{n}(X_i - \mu)\right)^4 = \sum_{i=1}^{n}(X_i - \mu)^4 + \sum_{i}\sum_{j>i} \frac{4!}{2!2!}(X_i - \mu)^2(X_j - \mu)^2 + T$$

Where $T$ represent a number of terms such that every term in it contains a factor like $(X_i - \mu)$
Note that $E[(X_i - \mu)(X_j - \mu)^3] = 0$ etc. because $X_i$ are independent.

- Let us assume $X_i$ have finite fourth moment

$$\left(\sum_{i=1}^{n}(X_i-\mu)\right)^4 = \sum_{i=1}^{n}(X_i-\mu)^4+\sum_{i}\sum_{j>i}\ \frac{4!}{2!2!}\ (X_i-\mu)^2(X_j-\mu)^2+T$$

  Where $T$ represent a number of terms such that every term in it contains a factor like $(X_i-\mu)$
  Note that $E[(X_i-\mu)(X_j-\mu)^3]=0$ etc. because $X_i$ are independent.

- Hence we get

$$E\left[\left(\sum_{i=1}^{n}(X_i-\mu)\right)^4\right] = nE[(X_i-\mu)^4]+3n(n-1)\sigma^4 \le C'n^2$$

- Now we can get, using Markov inequality

- Now we can get, using Markov inequality

$$P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] \;=\; P\left[|S_n - n\mu| > n\epsilon\right]$$

- Now we can get, using Markov inequality

$$\begin{aligned} P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] &= P\left[|S_n - n\mu| > n\epsilon\right] \\ &= P\left[\left|\sum_{i=1}^{n}(X_i - \mu)\right| > n\epsilon\right] \end{aligned}$$

- Now we can get, using Markov inequality

$$
\begin{aligned}
P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] &= P\left[|S_n - n\mu| > n\epsilon\right] \\
&= P\left[\left|\sum_{i=1}^{n}(X_i - \mu)\right| > n\epsilon\right] \\
&\leq \frac{E\left(\sum_{i=1}^{n}(X_i - \mu)\right)^4}{(n\epsilon)^4}
\end{aligned}
$$

- Now we can get, using Markov inequality

$$
\begin{aligned}
P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] &= P\left[|S_n - n\mu| > n\epsilon\right] \\
&= P\left[\left|\sum_{i=1}^{n}(X_i - \mu)\right| > n\epsilon\right] \\
&\leq \frac{E\left(\sum_{i=1}^{n}(X_i - \mu)\right)^4}{(n\epsilon)^4} \\
&\leq \frac{C' n^2}{n^4 \epsilon^4} = \frac{C}{n^2}
\end{aligned}
$$

- Now we can get, using Markov inequality

$$\begin{aligned} P\left[\left|\frac{S_n}{n} - \mu\right| > \epsilon\right] &= P\left[|S_n - n\mu| > n\epsilon\right] \\ &= P\left[\left|\sum_{i=1}^{n}(X_i - \mu)\right| > n\epsilon\right] \\ &\leq \frac{E\left(\sum_{i=1}^{n}(X_i - \mu)\right)^4}{(n\epsilon)^4} \\ &\leq \frac{C' n^2}{n^4 \epsilon^4} = \frac{C}{n^2} \end{aligned}$$

- Since $\sum_n \frac{C}{n^2} < \infty$, we get $\frac{S_n}{n} \stackrel{a.s.}{\to} \mu$

- Strong law of large numbers says

$$\frac{S_n}{n} \stackrel{a.s.}{\rightarrow} \mu \quad \text{where } S_n = \sum_{i=1}^{n} X_i, \quad X_i \text{ iid}, \quad EX_i = \mu$$

- Strong law of large numbers says

$$\frac{S_n}{n} \stackrel{a.s.}{\rightarrow} \mu \quad \text{where } S_n = \sum_{i=1}^{n} X_i, \quad X_i \text{ iid}, \quad EX_i = \mu$$

- We proved it assuming finite fourth moment of $X_i$.

- Strong law of large numbers says

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu \quad \text{where } \; S_n = \sum_{i=1}^{n} X_i, \;\; X_i \text{ iid}, \;\; EX_i = \mu$$

- We proved it assuming finite fourth moment of $X_i$.
- This is only for illustration

- Strong law of large numbers says

$$\frac{S_n}{n} \stackrel{a.s.}{\rightarrow} \mu \quad \text{where } S_n = \sum_{i=1}^{n} X_i, \; X_i \text{ iid}, \; EX_i = \mu$$

- We proved it assuming finite fourth moment of $X_i$.
- This is only for illustration
- Strong law holds without any such assumptions on moments

- Strong law of large numbers says

$$\frac{S_n}{n} \stackrel{a.s.}{\rightarrow} \mu \quad \text{ where } \; S_n = \sum_{i=1}^{n} X_i, \;\; X_i \text{ iid}, \;\; EX_i = \mu$$

- We proved it assuming finite fourth moment of $X_i$.
- This is only for illustration
- Strong law holds without any such assumptions on moments
- Strong law of large numbers says that sample mean converges to the expectation with probability one.

# Convergence in probability Vs Almost sure convergence

- $X_n \xrightarrow{P} X_0$
  $\Leftrightarrow \lim_{n\to\infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$

# Convergence in probability Vs Almost sure convergence

- $X_n \xrightarrow{P} X_0$
  $\Leftrightarrow \lim_{n\to\infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$
- $X_n \xrightarrow{a.s.} X_0$
  $\Leftrightarrow \lim_{n\to\infty} \ P\left(\cup_{k=n}^{\infty} \ [|X_k - X_0| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$

# Convergence in probability Vs Almost sure convergence

- $X_n \xrightarrow{P} X_0$
  $\Leftrightarrow \lim_{n\to\infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$
- $X_n \xrightarrow{a.s.} X_0$
  $\Leftrightarrow \lim_{n\to\infty} \ P\left(\cup_{k=n}^{\infty} \ [|X_k - X_0| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$
- One can intuitively see why convergence almost surely is a much stronger notion of convergence.

# Example

- $\Omega = [0, 1]$. Sequence of binary random variables: $X_{nk},\ k = 1, \cdots, n, n = 1, 2, \cdots$, defined by

$$X_{nk}(\omega) = 1 \ \text{ iff } \ \frac{k-1}{n} \leq \omega < \frac{k}{n},\ 1 \leq k \leq n, n = 1, 2, \cdots$$

That is, the sequence is $X_{11}, X_{21}, X_{22}, X_{31}, X_{32}, X_{33}, \cdots$.

# Example

- $\Omega = [0, 1]$. Sequence of binary random variables: $X_{nk},\ k = 1, \cdots, n, n = 1, 2, \cdots$, defined by
$$X_{nk}(\omega) = 1 \ \text{ iff } \ \frac{k-1}{n} \leq \omega < \frac{k}{n},\ 1 \leq k \leq n, n = 1, 2, \cdots$$
That is, the sequence is $X_{11}, X_{21}, X_{22}, X_{31}, X_{32}, X_{33}, \cdots$.
- One can show the sequence converges to zero in probability.

# Example

- $\Omega = [0, 1]$. Sequence of binary random variables: $X_{nk},\ k = 1, \cdots, n, n = 1, 2, \cdots$, defined by

$$X_{nk}(\omega) = 1 \ \text{ iff } \ \frac{k-1}{n} \leq \omega < \frac{k}{n},\ 1 \leq k \leq n, n = 1, 2, \cdots$$

  That is, the sequence is $X_{11}, X_{21}, X_{22}, X_{31}, X_{32}, X_{33}, \cdots$.
- One can show the sequence converges to zero in probability.
- But, $P[X_{nk} \rightarrow 0] = 0$!

$$X_{nk}(\omega) = 1 \text{ iff } \frac{k-1}{n} \leq \omega < \frac{k}{n},\ 1 \leq k \leq n$$

# Convergence in $r^{th}$ mean

- We say that a sequence $X_n$ converges in $r^{th}$ mean to $X$ if $E[|X_n|^r] < \infty, \ \forall n, \ E[|X|^r] < \infty$ and

$$E\left[|X_n - X|^r\right] \to 0 \quad \text{as} \quad n \to \infty$$

# Convergence in $r^{th}$ mean

- We say that a sequence $X_n$ converges in $r^{th}$ mean to $X$ if $E[|X_n|^r] < \infty, \ \forall n, \ E[|X|^r] < \infty$ and

$$E\left[|X_n - X|^r\right] \to 0 \quad \text{as} \quad n \to \infty$$

- Denoted as $X_n \xrightarrow{r} X$

# Convergence in $r^{th}$ mean

- We say that a sequence $X_n$ converges in $r^{th}$ mean to $X$ if $E[|X_n|^r] < \infty, \ \forall n, \ E[|X|^r] < \infty$ and

$$E\left[|X_n - X|^r\right] \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty$$

- Denoted as $X_n \xrightarrow{r} X$
- Consider our old example of binary random variables

$$P[X_n = 1] = \frac{1}{n} \qquad P[X_n = 0] = 1 - \frac{1}{n}$$

# Convergence in $r^{th}$ mean

- We say that a sequence $X_n$ converges in $r^{th}$ mean to $X$ if $E[|X_n|^r] < \infty,\ \forall n,\ E[|X|^r] < \infty$ and

$$E\left[|X_n - X|^r\right] \rightarrow 0 \qquad \text{as} \quad n \rightarrow \infty$$

- Denoted as $X_n \xrightarrow{r} X$
- Consider our old example of binary random variables

$$P[X_n = 1] = \frac{1}{n} \qquad P[X_n = 0] = 1 - \frac{1}{n}$$

- All moments of $X_n$ are finite and we have

# Convergence in $r^{th}$ mean

- We say that a sequence $X_n$ converges in $r^{th}$ mean to $X$ if $E[|X_n|^r] < \infty,\ \forall n,\ E[|X|^r] < \infty$ and

$$E\left[|X_n - X|^r\right] \to 0 \qquad \text{as} \quad n \to \infty$$

- Denoted as $X_n \xrightarrow{r} X$
- Consider our old example of binary random variables

$$P[X_n = 1] = \frac{1}{n} \qquad P[X_n = 0] = 1 - \frac{1}{n}$$

- All moments of $X_n$ are finite and we have

$$E\left[|X_n - 0|^2\right] = \frac{1}{n} \to 0$$

# Convergence in $r^{th}$ mean

- We say that a sequence $X_n$ converges in $r^{th}$ mean to $X$ if $E[|X_n|^r] < \infty, \ \forall n, \ E[|X|^r] < \infty$ and
$$E\left[|X_n - X|^r\right] \to 0 \quad \text{as} \quad n \to \infty$$
- Denoted as $X_n \xrightarrow{r} X$
- Consider our old example of binary random variables
$$P[X_n = 1] = \frac{1}{n} \qquad P[X_n = 0] = 1 - \frac{1}{n}$$
- All moments of $X_n$ are finite and we have
$$E\left[|X_n - 0|^2\right] = \frac{1}{n} \to 0$$
- Hence $X_n \xrightarrow{2} 0$.

# Convergence in $r^{th}$ mean

- We say that a sequence $X_n$ converges in $r^{th}$ mean to $X$ if $E[|X_n|^r] < \infty,\ \forall n,\ E[|X|^r] < \infty$ and

$$E\left[|X_n - X|^r\right] \to 0 \quad \text{as} \quad n \to \infty$$

- Denoted as $X_n \xrightarrow{r} X$
- Consider our old example of binary random variables

$$P[X_n = 1] = \frac{1}{n} \qquad P[X_n = 0] = 1 - \frac{1}{n}$$

- All moments of $X_n$ are finite and we have

$$E\left[|X_n - 0|^2\right] = \frac{1}{n} \to 0$$

- Hence $X_n \xrightarrow{2} 0$.
- In this example $X_n$ converges in $r^{th}$ mean for all $r$

- Suppose $X_n \stackrel{r}{\to} X$. Then, by Markov inequality

$$P\left[|X_n - X| > \epsilon\right] \leq \frac{E\left[|X_n - X|^r\right]}{\epsilon^r} \rightarrow 0$$

- Suppose $X_n \xrightarrow{r} X$. Then, by Markov inequality

$$P\left[|X_n - X| > \epsilon\right] \leq \frac{E\left[|X_n - X|^r\right]}{\epsilon^r} \rightarrow 0$$

- Hence

$$X_n \xrightarrow{r} X \quad \Rightarrow \quad X_n \xrightarrow{P} X$$

- Suppose $X_n \stackrel{r}{\to} X$. Then, by Markov inequality

$$P\left[|X_n - X| > \epsilon\right] \leq \frac{E\left[|X_n - X|^r\right]}{\epsilon^r} \to 0$$

- Hence

$$X_n \stackrel{r}{\to} X \quad \Rightarrow \quad X_n \stackrel{P}{\to} X$$

- Consider the sequence $X_n$ with

$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = e^n] = \frac{1}{n}$$

- Suppose $X_n \xrightarrow{r} X$. Then, by Markov inequality

$$P\left[|X_n - X| > \epsilon\right] \leq \frac{E\left[|X_n - X|^r\right]}{\epsilon^r} \to 0$$

- Hence

$$X_n \xrightarrow{r} X \quad \Rightarrow \quad X_n \xrightarrow{P} X$$

- Consider the sequence $X_n$ with

$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = e^n] = \frac{1}{n}$$

We have $P[|X_n - 0| > \epsilon] = \frac{1}{n} \;\Rightarrow\; X_n \xrightarrow{P} 0$

- Suppose $X_n \xrightarrow{r} X$. Then, by Markov inequality

$$P\left[|X_n - X| > \epsilon\right] \leq \frac{E\left[|X_n - X|^r\right]}{\epsilon^r} \to 0$$

- Hence

$$X_n \xrightarrow{r} X \quad \Rightarrow \quad X_n \xrightarrow{P} X$$

- Consider the sequence $X_n$ with

$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = e^n] = \frac{1}{n}$$

We have $P[|X_n - 0| > \epsilon] = \frac{1}{n} \;\Rightarrow\; X_n \xrightarrow{P} 0$

- But, $E|X_n - 0|^r = \frac{e^{rn}}{n}$ and hence $X_n$ does not converge in $r^{th}$ mean.

- In general, neither of convergence almost surely and in $r^{th}$ mean imply the other.

- In general, neither of convergence almost surely and in $r^{th}$ mean imply the other.
- We can generate counter examples for this easily.

- In general, neither of convergence almost surely and in $r^{th}$ mean imply the other.
- We can generate counter examples for this easily.
- However, if all $X_n$ take values in a bounded interval, then almost sure convergence implies $r^{th}$ mean convergence

- Consider sequence $X_n$ where $X_n$ are independent with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- Consider sequence $X_n$ where $X_n$ are independent with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- Consider sequence $X_n$ where $X_n$ are independent with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- Assume $a_n \to 0$ so that $X_n \xrightarrow{P} 0$

- Consider sequence $X_n$ where $X_n$ are independent with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- Assume $a_n \to 0$ so that $X_n \stackrel{P}{\to} 0$
- By Borel-Cantelli lemma

$$X_n \stackrel{a.s.}{\to} 0 \quad \Leftrightarrow \quad \sum_n a_n < \infty$$

- Consider sequence $X_n$ where $X_n$ are independent with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- Assume $a_n \to 0$ so that $X_n \xrightarrow{P} 0$
- By Borel-Cantelli lemma

$$X_n \xrightarrow{a.s.} 0 \quad \Leftrightarrow \quad \sum_n a_n < \infty$$

- For convergence in $r^{th}$ mean we need

$$E\left[|X_n - 0|^r\right] = (c_n)^r \ a_n \ \rightarrow \ 0$$

- Consider sequence $X_n$ where $X_n$ are independent with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- Assume $a_n \to 0$ so that $X_n \xrightarrow{P} 0$
- By Borel-Cantelli lemma

$$X_n \xrightarrow{a.s.} 0 \quad \Leftrightarrow \quad \sum_n a_n < \infty$$

- For convergence in $r^{th}$ mean we need

$$E\left[|X_n - 0|^r\right] = (c_n)^r \ a_n \ \to \ 0$$

- Take $a_n = \frac{1}{n}$ and $c_n = 1$. Then $X_n \xrightarrow{r} 0$ but the sequence does not converge almost surely.

- Consider sequence $X_n$ where $X_n$ are independent with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- Assume $a_n \to 0$ so that $X_n \stackrel{P}{\to} 0$
- By Borel-Cantelli lemma

$$X_n \stackrel{a.s.}{\to} 0 \quad \Leftrightarrow \quad \sum_n a_n < \infty$$

- For convergence in $r^{th}$ mean we need

$$E\left[|X_n - 0|^r\right] = (c_n)^r \ a_n \ \to \ 0$$

- Take $a_n = \frac{1}{n}$ and $c_n = 1$. Then $X_n \stackrel{r}{\to} 0$ but the sequence does not converge almost surely.
- Take $a_n = \frac{1}{n^2}$ and $c_n = e^n$. Then $X_n \stackrel{a.s.}{\to} 0$ but the sequence does not converge in $r^{th}$ mean for any $r$.

- Let $X_n \xrightarrow{r} X$. Then

- Let $X_n \xrightarrow{r} X$. Then
  1. $E[|X_n|^r] \rightarrow E[|X|^r]$

- Let $X_n \xrightarrow{r} X$. Then
  1. $E[|X_n|^r] \rightarrow E[|X|^r]$
  2. $X_n \xrightarrow{s} X, \ \forall s < r$

- Let $X_n \xrightarrow{r} X$. Then
  1. $E[|X_n|^r] \rightarrow E[|X|^r]$
  2. $X_n \xrightarrow{s} X, \ \forall s < r$
- The proofs are straight-forward but we omit the proofs

# Convergence in distribution

- Let $F_n$ be the df of $X_n$, $n = 1, 2, \cdots$. Let $X$ be a rv with df $F$.

# Convergence in distribution

- Let $F_n$ be the df of $X_n$, $n = 1, 2, \cdots$. Let $X$ be a rv with df $F$.
- Sequence $X_n$ is said to converge to $X$ **in distribution** if

$$F_n(x) \to F(x), \;\; \forall x \;\text{ where } F \text{ is continuous}$$

# Convergence in distribution

- Let $F_n$ be the df of $X_n$, $n = 1, 2, \cdots$. Let $X$ be a rv with df $F$.
- Sequence $X_n$ is said to converge to $X$ **in distribution** if

$$F_n(x) \to F(x), \ \ \forall x \ \ \text{where } F \text{ is continuous}$$

- We denote this as

$$X_n \stackrel{d}{\to} X, \quad \text{or} \quad X_n \stackrel{L}{\to} X, \quad \text{or } F_n \stackrel{w}{\to} F$$

# Convergence in distribution

- Let $F_n$ be the df of $X_n$, $n = 1, 2, \cdots$. Let $X$ be a rv with df $F$.
- Sequence $X_n$ is said to converge to $X$ **in distribution** if

$$F_n(x) \to F(x), \;\; \forall x \;\; \text{where } F \text{ is continuous}$$

- We denote this as

$$X_n \xrightarrow{d} X, \quad \text{or} \quad X_n \xrightarrow{L} X, \quad \text{or } F_n \xrightarrow{w} F$$

- This is also known as **convergence in law** or weak convergence

# Convergence in distribution

- Let $F_n$ be the df of $X_n$, $n = 1, 2, \cdots$. Let $X$ be a rv with df $F$.
- Sequence $X_n$ is said to converge to $X$ **in distribution** if

$$F_n(x) \to F(x), \quad \forall x \quad \text{where } F \text{ is continuous}$$

- We denote this as

$$X_n \xrightarrow{d} X, \quad \text{or} \quad X_n \xrightarrow{L} X, \quad \text{or } F_n \xrightarrow{w} F$$

- This is also known as **convergence in law** or weak convergence
- Note that here we are essentially talking about convergence of distribution functions.

# Convergence in distribution

- Let $F_n$ be the df of $X_n$, $n = 1, 2, \cdots$. Let $X$ be a rv with df $F$.
- Sequence $X_n$ is said to converge to $X$ **in distribution** if

$$F_n(x) \to F(x), \;\; \forall x \;\; \text{where } F \text{ is continuous}$$

- We denote this as

$$X_n \xrightarrow{d} X, \quad \text{or} \quad X_n \xrightarrow{L} X, \quad \text{or } F_n \xrightarrow{w} F$$

- This is also known as **convergence in law** or weak convergence
- Note that here we are essentially talking about convergence of distribution functions.
- Convergence in probability implies convergence in distribution

# Convergence in distribution

- Let $F_n$ be the df of $X_n$, $n = 1, 2, \cdots$. Let $X$ be a rv with df $F$.
- Sequence $X_n$ is said to converge to $X$ **in distribution** if

$$F_n(x) \to F(x), \;\; \forall x \;\; \text{where } F \text{ is continuous}$$

- We denote this as

$$X_n \xrightarrow{d} X, \quad \text{or} \quad X_n \xrightarrow{L} X, \quad \text{or } F_n \xrightarrow{w} F$$

- This is also known as **convergence in law** or weak convergence
- Note that here we are essentially talking about convergence of distribution functions.
- Convergence in probability implies convergence in distribution
- The converse is not true. (e.g., sequence of iid random variables)

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

$$P[N_n > a] = (P[X_i > a])^n$$

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

$$P[N_n > a] = (P[X_i > a])^n = (1-a)^n, \quad 0 < a < 1$$

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

$$P[N_n > a] = (P[X_i > a])^n = (1-a)^n, \quad 0 < a < 1$$

$$P[Y_n > y] = P[N_n > y/n]$$

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

$$P[N_n > a] = (P[X_i > a])^n \ = (1-a)^n, \ \ 0 < a < 1$$

$$P[Y_n > y] = P[N_n > y/n] = \left(1 - \frac{y}{n}\right)^n, \quad \text{if} \ \ n > y$$

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

$$P[N_n > a] = \left(P[X_i > a]\right)^n = (1-a)^n, \quad 0 < a < 1$$

$$P[Y_n > y] = P[N_n > y/n] = \left(1 - \frac{y}{n}\right)^n, \quad \text{if } \ n > y$$

- Hence for any $y$

$$\lim_{n\to\infty} P[Y_n > y] = \lim_{n\to\infty} \left(1 - \frac{y}{n}\right)^n$$

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

$$P[N_n > a] = (P[X_i > a])^n = (1-a)^n, \quad 0 < a < 1$$

$$P[Y_n > y] = P[N_n > y/n] = \left(1 - \frac{y}{n}\right)^n, \quad \text{if } \ n > y$$

- Hence for any $y$

$$\lim_{n\to\infty} P[Y_n > y] = \lim_{n\to\infty} \left(1 - \frac{y}{n}\right)^n = e^{-y}$$

# Examples

- $X_1, X_2, \cdots$ be iid; uniform over $(0,\ 1)$
- $N_n = \min(X_1, \cdots, X_n)$, $Y_n = nN_n$.
  Does $Y_n$ converge in distribution?

$$P[N_n > a] = (P[X_i > a])^n \ = (1-a)^n, \ \ 0 < a < 1$$

$$P[Y_n > y] = P[N_n > y/n] = \left(1 - \frac{y}{n}\right)^n, \quad \text{if} \ \ n > y$$

- Hence for any $y$

$$\lim_{n\to\infty} P[Y_n > y] = \lim_{n\to\infty} \left(1 - \frac{y}{n}\right)^n \ = e^{-y}$$

- The sequence converges in distribution to an exponential rv

- $X_n \xrightarrow{d} X$
  $\Leftrightarrow \quad F_n(x) \rightarrow F(x), \quad \forall x$ where $F$ is continuous

- $X_n \xrightarrow{d} X$
  $\Leftrightarrow \quad F_n(x) \rightarrow F(x), \quad \forall x$ where $F$ is continuous
- This means that the sequence of functions $F_n$ converge point-wise and the limit function is a distribution function.

- $X_n \xrightarrow{d} X$
  $\Leftrightarrow \quad F_n(x) \rightarrow F(x), \quad \forall x$ where $F$ is continuous
- This means that the sequence of functions $F_n$ converge point-wise and the limit function is a distribution function.
- In general, $X_n \xrightarrow{d} X$ does not imply that the pdf's or pmf's converge point-wise to the limit pdf or pmf.

- $X_n \xrightarrow{d} X$
  $\Leftrightarrow \quad F_n(x) \rightarrow F(x), \quad \forall x$ where $F$ is continuous
- This means that the sequence of functions $F_n$ converge point-wise and the limit function is a distribution function.
- In general, $X_n \xrightarrow{d} X$ does not imply that the pdf's or pmf's converge point-wise to the limit pdf or pmf.
- However if the sequence of pmf's (or pdf's) converge point-wise and the limit is a pmf (or pdf) then we have $X_n \xrightarrow{d} X$.

- $X_n \xrightarrow{d} X$
  $\Leftrightarrow \quad F_n(x) \to F(x), \quad \forall x$ where $F$ is continuous
- This means that the sequence of functions $F_n$ converge point-wise and the limit function is a distribution function.
- In general, $X_n \xrightarrow{d} X$ does not imply that the pdf's or pmf's converge point-wise to the limit pdf or pmf.
- However if the sequence of pmf's (or pdf's) converge point-wise and the limit is a pmf (or pdf) then we have $X_n \xrightarrow{d} X$.
- $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$

- $X_n \xrightarrow{d} X$
  $\Leftrightarrow \quad F_n(x) \to F(x), \quad \forall x$ where $F$ is continuous
- This means that the sequence of functions $F_n$ converge point-wise and the limit function is a distribution function.
- In general, $X_n \xrightarrow{d} X$ does not imply that the pdf's or pmf's converge point-wise to the limit pdf or pmf.
- However if the sequence of pmf's (or pdf's) converge point-wise and the limit is a pmf (or pdf) then we have $X_n \xrightarrow{d} X$.
- $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$
- $X_n \xrightarrow{d} k \Rightarrow X_n \xrightarrow{P} k$, where $k$ is a constant