Prop: Let $Z_1, Z_2, Z_3, \ldots$ be random variables. Suppose that $Z_n \to Z$ almost surely, i.e. $P\left(\lim_{n \to \infty} Z_n = Z\right) = 1$. Then $Z_n \to Z$ in probability, i.e. for any $\varepsilon > 0$, $P(|Z_n - Z| > \varepsilon) \to 0$ as $n \to \infty$.

PF: Suppose $Z_n \overset{a.s.}{\to} Z$. Then the set of points

$\{\omega \in \Omega : \lim_{n \to \infty} Z_n(\omega) = Z(\omega)\}$ has measure 1, and hence

$\{\omega \in \Omega : \lim_{n \to \infty} Z_n(\omega) \neq Z(\omega)\}$ has measure 0. Denote the latter set by $\Theta$. Fix $\varepsilon > 0$ and let

$A_n = \{\omega \in \Omega : \exists m > n, |Z_m - Z| > \varepsilon\}.$

Then $\{A_n\}$ is a decreasing sequence of events:

$A_n \supseteq A_{n+1} \supseteq \ldots$ and it decreases to the set

$A_\infty = \bigcap_{n=1}^{\infty} A_n.$

Now any point $\omega \in \Theta^c$ is s.t.

$\lim_{n \to \infty} Z_n(\omega) = Z(\omega) \quad \Rightarrow \quad |Z_n(\omega) - Z(\omega)| < \varepsilon \quad \forall n \geq N$

for some $N$. Thus, for all $n \geq N$, $\omega \notin A_n$ and hence $\omega \notin A_\infty$. [In other words, if $\omega \in A_\infty$ then $Z_n(\omega) \not\to Z(\omega)$ as $n \to \infty$.]

Hence, $P(A_\infty) = P\left(\bigcap_{n=1}^{\infty} A_n\right) \leq P(\Theta) = 0$.

i.e. $P(Z_n \not\to Z) = 0$.
By continuity of probabilities,

\[ P(An) \to P(\bigcap_{n=1}^{\infty} A_n) = 0 \quad \text{as } n \to \infty. \]

Therefore,

\[ P(\left| z_n - z \right| \geq \varepsilon) \leq P(An) \to 0 \quad \text{as } n \to \infty, \]

and we have convergence in probability.
Last time: Different notions of convergence:
- pointwise
- almost surely
- in probability

Convergence a.s. \( \Rightarrow \) convergence in probability

But converse is not true

**Example 1:** Let \( \{ Z_n \} \) be independent RVs, with
\[
P(Z_n = 1) = \frac{1}{n} = 1 - P(Z_n = 0).
\]

Then \( Z_n \xrightarrow{p} 0 \) (in probability)

since
\[
\lim_{n \to \infty} P(|Z_n - 0| \geq \varepsilon) = \lim_{n \to \infty} P(Z_n = 1)
\]
\[
= \lim_{n \to \infty} \frac{1}{n} = 0.
\]

On the other hand, by the Borel-Cantelli lemma,
\[
P(Z_n = 1 \text{ i.o.}) = 1, \text{ so } P(Z_n \to 0) = 0,
\]

and hence \( Z_n \) does NOT converge to 0 a.s.
(for \( Z_n \xrightarrow{a.s.} 0 \) we would need \( P(Z_n \to 0) = 1 \)).
More details:
Since \( \sum_{n=1}^{\infty} P(Z_n = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty \) and the events \( \{Z_n = 1\} \) are independent, then BC lemma \( \Rightarrow P(\limsup_{n} \{Z_n = 1\}) = 1 \).
Hence, \( Z_n \not\to 0 \) a.s. In fact, the set on which it doesn't converge to 0 has probability 1.

**Laws of Large Numbers (LLNs)**

**Thm [Weak LLNs]:** Let \( X_1, X_2, \ldots \) be a sequence of independent random variables, each having the same mean \( \mu \) and each having variance \( \text{Var}(X_i) \leq \sigma^2 < \infty \). Then \( \forall \varepsilon > 0 \),

\[
\lim_{n \to \infty} P \left( \frac{1}{n} (X_1 + X_2 + \cdots + X_n) - \mu \geq \varepsilon \right) = 0.
\]

i.e. If we let \( S_n = \frac{1}{n} \sum_{i=1}^{n} X_i \), then \( \frac{S_n}{n} \to \mu \) as \( n \to \infty \).

**Pf:** Let \( S_n = \frac{1}{n} \sum_{i=1}^{n} X_i \). Then \( E[S_n] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] \) by linearity

\[
\Rightarrow E\left[ \frac{S_n}{n} \right] = \frac{1}{n} \cdot n \cdot \mu = \mu.
\]

Also, \( \text{Var} \left( \frac{S_n}{n} \right) = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}(X_i) \)

\[
\leq \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n} \quad \text{since we assumed } \text{Var}(X_i) \leq \sigma^2 \quad \forall i.\]
Then by Chebyshev’s Ineq. we have

\[ P \left( \left| \frac{1}{n} (X_1 + \ldots + X_n) - \mu \right| \geq \varepsilon \right) \leq \frac{\text{var}(Y)}{\varepsilon^2 n} \rightarrow 0 \text{ as } n \rightarrow \infty \]

as required.

Version 1
Thm [**Strong LLN**] : Let \( X_1, X_2, \ldots \) be a sequence of independent random variables, each having the same mean \( \mu \) and each having \( E[(X_i - \mu)^4] \leq a < \infty \). Then

\[ P \left( \lim_{n \rightarrow \infty} \frac{1}{n} (X_1 + \ldots + X_n) = \mu \right) = 1, \]

i.e. the partial averages \( \frac{S_n}{n} \rightarrow \mu \) almost surely.

**Differences between Weak & Strong LLNs?**

**Weak** : assumes finite variance
convergence in probability

**Strong** : assumes finite 4\textsuperscript{th} moment (version 1)
convergence almost surely

\[ \downarrow \]

develop another version which only requires the mean to be finite, but as a penalty requires RVs to be i.i.d.
instead of merely independent.
Eliminating the moment conditions

def: A collection of random variables \( \{X_\alpha \}_{\alpha \in I} \) are identically distributed if for any Borel measurable function \( f: \mathbb{R} \to \mathbb{R} \), the expected value \( \mathbb{E}[f(X_\alpha)] \) does not depend on \( \alpha \), i.e. is the same for all \( \alpha \in I \).

Remark: It follows that \( \{X_\alpha \}_{\alpha \in I} \) are identically distributed iff \( \forall x \in \mathbb{R}, \ P(X_\alpha \leq x) \) does not depend on \( \alpha \).

def: A collection of RVs \( \{X_\alpha \}_{\alpha \in I} \) are i.i.d. if they are independent and identically distributed.

Thm [Strong LLNs Version 2]: Let \( X_1, X_2, \ldots \) be a sequence of i.i.d. random variables, each having finite mean \( \mu \). Then

\[
P \left( \lim_{n \to \infty} \frac{1}{n} (X_1 + \cdots + X_n) = \mu \right) = 1.
\]

i.e. \( S_n \xrightarrow{a.s.} \mu \) where \( S_n = X_1 + \cdots + X_n \)