Almost sure convergence

X_n converges to X almost surely if

$$P\left(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1.$$

- The set of outcomes where the convergence does not happen has measure 0. $P\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\} = 0$.
- Consider $\Omega = [0, 1]$ where you pick a number uniformly in $[0, 1]$. Let $X_n(\omega) = \omega^n$ for all $\omega \in \Omega$ and $X(\omega) = 0$ for all ω.
- $X_n(\omega) \to X(\omega)$ for $\omega \in [0, 1)$.
- $X_n(\omega) \nrightarrow X(\omega)$ for $\omega = 1$ and $\mathbb{P}\{\omega = 1\}$.
- This is almost sure convergence as $\mathbb{P}\{[0, 1)\} = 1$.

Almost sure (a.s.) convergence

X_n converges to X almost surely if

$$P\left(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1.$$

- Example 2: Strong law of large numbers (SLLN).

  Let $\{X_n, n \geq 0\}$ denote a sequence of i.i.d random variables with mean μ and denote $S_n = \sum_{i=1}^{n} X_i$. Then $\frac{S_n}{n} \to \mu$ a.s.

- Toss a biased coin (probability of head is μ) repeatedly. What is ω and Ω?
- Let X_i denote the outcome of the i^{th} toss and S_n denotes the number of heads in n tosses.
- The empirical mean is given by $\frac{S_n}{n}$.

Detour: Incremental formula for sample mean

- Now that we know $\frac{S_n}{n} \to \mu$ we can use $\hat{\mu_n} := \frac{S_n}{n}$ as an 'estimator' for the mean especially in cases when the underlying distribution is not known.
- Note that the estimator $\hat{\mu_n}$ is a random variable. What is its cdf? what is its mean & Variance?
- $\hat{\mu_n} = \frac{S_n}{n}$ is an 'unbiased estimator' since $E[\hat{\mu_n}] = \mu$.
- $Var(\hat{\mu_n}) = \frac{\sigma^2}{n}$
- We will soon see CLT that will tell the CDF of $\hat{\mu_n}$ without any information on the law of X_i.

Detour: Incremental formula for sample mean

- Now given $\hat{\mu}_n$, suppose you see an additional sample X_{n+1}.
- How will you compute $\hat{\mu}_{n+1}$?
- Naive way : $\hat{\mu}_{n+1} = \frac{\sum_{i=1}^{n+1} X_i}{n+1}$.
- There is an incremental formula that uses $\hat{\mu}_n$.

$$\boxed{\hat{\mu}_{n+1} = \hat{\mu}_n + \frac{1}{n+1}\left[X_{n+1} - \hat{\mu}_n\right]}$$

- Such averaging formulas are used extensively in Reinforcement learning.

Recap: Modes of Convergence

$\{X_n, n \geq 0\}$ converges to X pointwise or surely if for all $\omega \in \Omega$ we have $\lim_{n\to\infty} X_n(\omega) = X(\omega)$

X_n converges to X almost surely if $P\left(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1.$

$\{X_n, n \geq 0\}$ is a sequence of i.i.d random variables with mean μ and $S_n = \sum_{i=1}^{n} X_i$. Then $\hat{\mu}_n := \frac{S_n}{n} \to \mu$ a.s. (SLLN)

- Estimator $\hat{\mu}_n$ has mean μ and Variance $\frac{\sigma^2}{n}$.
- $\hat{\mu}_{n+1} = \hat{\mu_n} + \frac{1}{n+1}\left[X_{n+1} - \hat{\mu_n}\right]$

Borel Cantelli Lemma

Self-Study: Theorem 7.5 (probabilitycourse.com)
Consider a sequence of random variables $X_1, X_2, \ldots$. If for all ϵ we have

$$\sum_{n=1}^{\infty} P(|X_n - X| > \epsilon) < \infty$$

then $X_n \to X$ a.s.

- This is only a sufficient condition for almost sure convergence!
- Thm 7.6 (HW) gives necessary and sufficient conditions.
- Lot of problems in probabilitycourse, practice them!

Another example of a.s. convergence

- Consider a uniform r.v. U and define $X_n = n1_{\{U \leq \frac{1}{n}\}}$.
- $X_n = n$ when $U \leq \frac{1}{n}$ and $X_n = 0$ otherwise.
- Given a realization of U, what can you say about the sequence $\{X_n\}$?
- Once an X_n is zero, all higher indexed variables are also zero!
- This happens for all realizations U other than $U = 0$. In this case since $0 \leq \frac{1}{n}$ for all n, $X_n's$ run off to infinity and we don't see convergence to 0.
- But $P(U = 0) = 0$.
- Does $E[X_n] \to 0$?
- Almost sure convergence does not imply their means converge!

Towards convergence in probability

- Now define $X_n = n1_{\{U_n \leq \frac{1}{n}\}}$ where $\{U_n\}$ are i.i.d uniform.
- $X_n = n$ when $U_n \leq \frac{1}{n}$ and $X_n = 0$ otherwise.
- What can you say about the sequence $\{X_n\}$?
- Is it true that once an X_n is zero, all higher indexed variables are also zero!? No!
- Every time (on every run of the experiment or every sample path), we will have a sequence of zero and non-zero values, where the non-zero values become rarer and rarer but will keep happening once in a while.
- On no sample path would you see convergence to zero but occurrence of non-zero values become rare.
- We now characterize this notion of convergence.

Convergence in probability (w.h.p)

X_n converges to X in probability if

$$\lim_{n\to\infty} P\left(|X_n - X| > \epsilon\right) = 0 \text{ for all } \epsilon > 0.$$

- How would you compute $P\left(|X_n - X| > \epsilon\right)$ when X_n, X are either continuous or discrete random variables ?
- Ex: $X_n = n$ with probability $\frac{1}{n}$ and $X_n = 0$ otherwise.
- $P(|X_n - X| > \epsilon) = P(X_n > \epsilon) = \frac{1}{n}$ when $n > \epsilon$.
- When $n < \epsilon$, we have $P(|X_n - X| > \epsilon) = 0$.
- Once $n > \epsilon$ we have $\lim_{n\to\infty} P(X_n > \epsilon) = \lim_{n\to\infty} \frac{1}{n} = 0$.
- X_n converges to 0 in probability, but not almost surely.
- a.s. convergence implies convergence in probability