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 $\frac{\sigma^2}{n}$.
- $\hat{\mu}_{n+1} = \hat{\mu_n} + \frac{1}{n+1}\left[X_{n+1} - \hat{\mu_n}\right]$

Section 5.1 on Markov and Chebyshev inequality is for self study

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.
- $P(|X_n - X| > \epsilon) = P(X_n = n) = \frac{1}{n}$ when $n > \epsilon$.
- When $n < \epsilon$, we have $P(|X_n - X| > \epsilon) = 0$.
- Once $n > \epsilon$ we have $\lim_{n\to\infty} P(X_n = n) = 0$.
- X_n converges to 0 in probability, but not almost surely.
- a.s. convergence implies convergence in probability

Convergence in r^{th} mean

X_n converges to X in r^{th} mean if

$$\lim_{n\to\infty} E[|X_n - X|^r] = 0.$$

- How will you compute $E[|X_n - X|^r]$?
- When $r = 2$, it is convergence in mean squared sense. In addition if $X = 0$, it implies that the second moments converge to 0.
- In the convergence in probability example, do we have convergence in mean or mean square?
- Convergence in r^{th} mean implies convergence in probability.

Weak convergence (in distribution)

X_n converges to X in distribution if

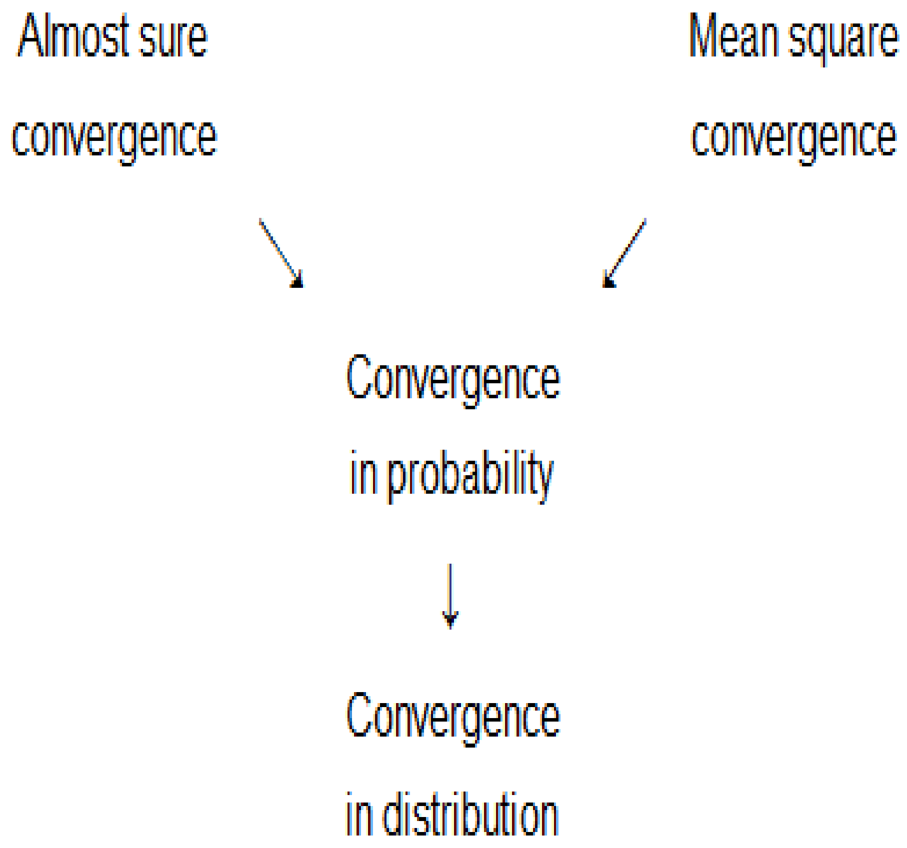
$$\boxed{\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \text{ for all continuity points of } F_X(\cdot).}$$

- a.s. convergence and convergence in probability imply convergence in distribution.
- Example: X_n is an exponential random variable with parameter λn.
- In this case, $F_{X_n}(x) = 1 - e^{-n\lambda x}$ and $F_X(x) = 1$ for all x.
- Note $x = 0$ is point of discontinuity as $F_X(0) = 1$ and $F_{X_n}(0) = 0$.
- HW EX2: X_n are i.i.d Binomial$(n, \frac{\lambda}{n})$. It converges in distribution to Poisson(λ).

Summary

Pointwise convergence	$\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for every ω
Almost sure convergence	$\lim_{n\to\infty} X_n(\omega) = X(\omega)$ almost surely
Convergence in probability	$\lim_{n\to\infty} \mathrm{P}(\lvert X_n - X\rvert > \epsilon) = 0$ for any $\varepsilon > 0$
Mean-square convergence	$\lim_{n\to\infty} \mathrm{E}[(X_n - X)^2] = 0$
Convergence in distribution	$\lim_{n\to\infty} F_n(x) = F(x)$ for any continuity point x

Relation between modes of convergence (no proofs)



https://en.wikipedia.org/wiki/Proofs_of_convergence_of_random_variables

Towards CLT

- Recall $\hat{\mu}_n = \frac{S_n}{n}$ where $S_n = \sum_{i=1}^{n} X_i$
- $\{X_i\}$ is i.i.d. with mean μ amnd variance σ^2.
- $E[\hat{\mu}_n] = \mu$ and $var(\hat{\mu}_n) = \frac{\sigma^2}{n}$
- Now consider $Y_n = \frac{\hat{\mu}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$. (centering and scaling). What is the mean and variance of Y_n?
- $E[Y_n] = 0$ and $Var(Y_n) = 1$. What is $F_{Y_n}(\cdot)$?
- What is $\lim_{n \to \infty} F_{Y_n}(\cdot)$? ANS: $\Phi(\cdot) = F_{N(0,1)}(\cdot)$
- In other words, Y_n converges to $Y = N(0,1)$ in distribution.

CLT

Let $\{X_n, n \geq 0\}$ denote a sequence of i.i.d random variables each with mean μ and variance $0 < \sigma^2 < \infty$.Denote $\hat{\mu}_n = \frac{\sum_{i=1}^{n} X_i}{n}$ and $Y_n = \frac{\hat{\mu}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$. Then Y_n converges to $N(0,1)$ in distribution.

- X_i could be ANY discrete or continuous r.v. with finite mean and variance.
- What is the consequence when $E[X_i] = 0$ and $Var(X_i) = 1$.
- In this case, $Y_n = \frac{S_n}{\sqrt{n}}$ and it converges in distribution to $N(0,1)$.
- $\frac{S_n}{n}$ converges almost surely to 0 but $\frac{S_n}{\sqrt{n}}$ converges to a random variable $\mathcal{N}(0,1)$.

CLT

Let $\{X_n, n \geq 0\}$ denote a sequence of i.i.d random variables each with mean μ and variance $0 < \sigma^2 < \infty$.Denote $\hat{\mu}_n = \frac{\sum_{i=1}^{n} X_i}{n}$ and $Y_n = \frac{\hat{\mu}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$. Then Y_n converges to $N(0, 1)$ in distribution.

- CLT given a way to find approximate disribution of $\hat{\mu}_n$.
- Note that for large enough n, we can use the approximation that $Y_n \sim \mathcal{N}(0, 1)$.
- Since Gaussianity is preserved under affine transformation, $\hat{\mu}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$

Example from probabilitycourse.com

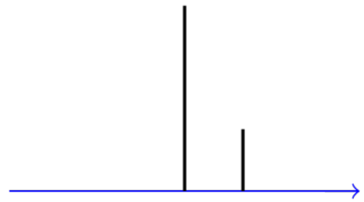
Assumptions:

- $X_1, X_2 \ldots$ are iid Bernoulli(p).
- $Z_n = \dfrac{X_1 + X_2 + \ldots + X_n - np}{\sqrt{np(1-p)}}$.

We choose $p = \frac{1}{3}$.

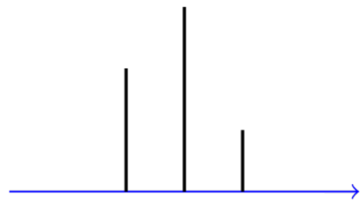
$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

PMF of Z_1



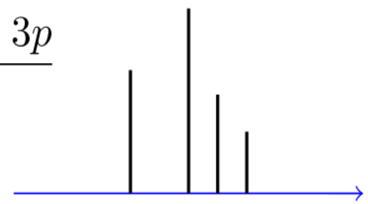
$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$

PMF of Z_2



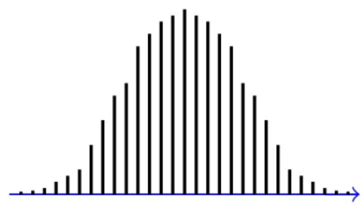
$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$

PMF of Z_3



$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - 30p}{\sqrt{30p(1-p)}}$$

PMF of Z_{30}



Normal Approximation based on CLT

- Let $S_n = X_1 + \ldots X_n$ where X_i are i.i.d. with mean μ and variance σ^2. If n is large, CDF of S_n can be approximated as follows.

$$P(S_n < c) \approx \Phi(z) \text{ where } z = \frac{c - n\mu}{\sigma\sqrt{n}}$$

Markov and Chebyschev inequalities (HW)

For a non-negative random variable X, we have for all $a > 0$

$$P(X > a) \leq \frac{E[X]}{a}$$

If X is a random variable with mean μ and variance σ^2, then for all $a > 0$,

$$P(|X - \mu| > a) \leq \frac{\sigma^2}{a^2}$$