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.

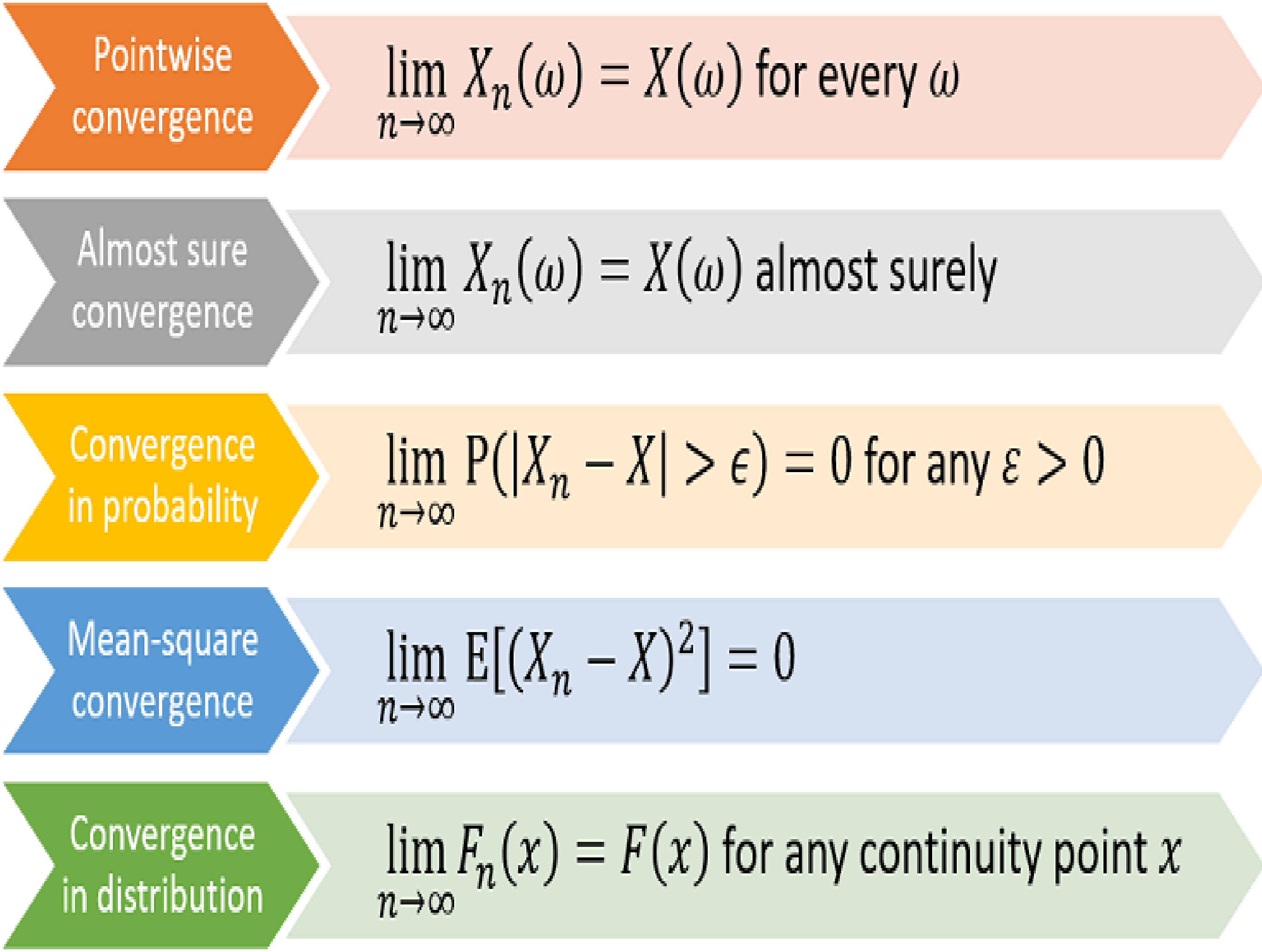
Weak convergence (in distribution)

X_n converges to X in distribution if

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

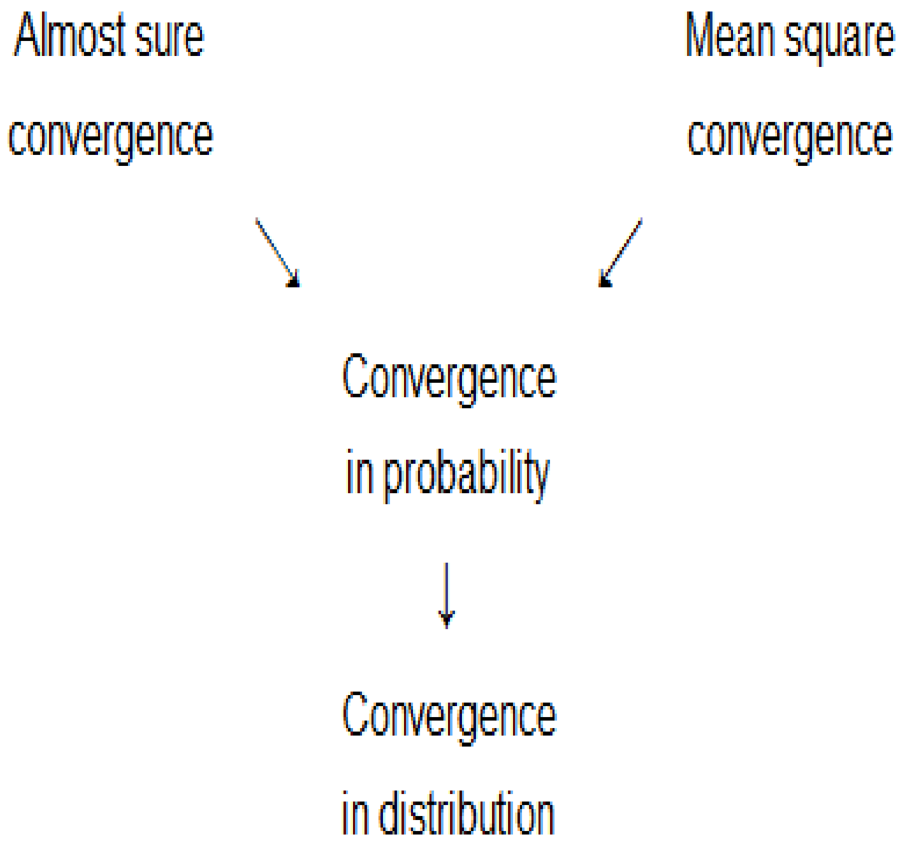
- a.s. convergence and convergence in probability imply convergence in distribution.
- Example: X_n is an exponential random variable with parameter λn.
- In this case, $F_{X_n}(x) = 1 - e^{-n\lambda x}$ and $F_X(x) = 1$ for all x.
- Note $x = 0$ is point of discontinuity as $F_X(0) = 1$ and $F_{X_n}(0)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

Summary



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

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

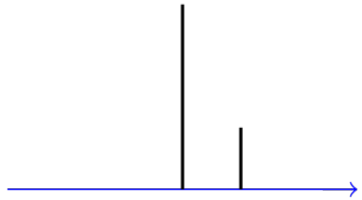
Assumptions:

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

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

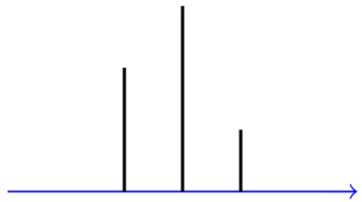
$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

PMF of Z_1



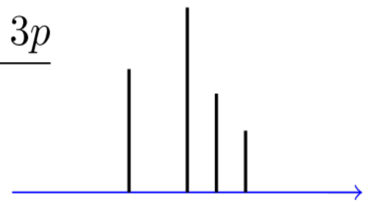
$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$

PMF of Z_2



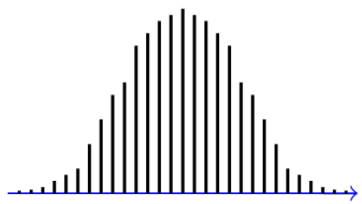
$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$

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.

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

https://www.youtube.com/watch?v=zeJD6dqJ5lo&t=111s