5.3 Convergence concepts

Given a sequence of Random variables $X_1, \ldots, X_n$ we want to study various notions of convergence to a random variable X:

(i) $X_n \xrightarrow{P} X$ Convergence in probability

(ii) $X_n \rightarrow X$ a.s. Convergence almost surely / with prob. 1.

(iii) $X_n \xrightarrow{d} X$ Convergence in distribution.

We already encountered (iii)

$X_n \xrightarrow{d} X$ if $F_{X_n}(x) \longrightarrow F_X(x)$ at all points x where F_X is continuous.

Recall Approximation of binomial distributions

if $X_n \sim$ Binomial (n, p_n)

and $np_n \to \lambda$ as $n \to \infty$

Then $X_n \xrightarrow{d} Y$ where $Y \sim$ Poisson (λ).

But, if $X_n \sim$ Binomial (n, p_n)

$p_n \to p$ as $n \to \infty$

Then $\frac{1}{\sqrt{n}}(X_n - np) \xrightarrow{d} Z$ where $Z \sim N(0, p(1-p))$

as a consequence of the central limit theorem to be learned in this chapter

Def. We say $X_n \to X$ in probability if

$\forall \varepsilon > 0,\quad P(|X_n - X| \geq \varepsilon) \to 0$ as $n \to \infty$.

Theorem Weak law of large numbers (WLLN)

Let $X_1, \dots, X_n$ are iid Random variables with $EX_i = \mu$, $\operatorname{var} X_i = \sigma^2 < \infty$

Define $\overline{X_n} = \frac{1}{n}(X_1 + \dots + X_n)$.

Then $\overline{X_n} \to \mu$ in probability.

Proof.

$$\forall \varepsilon > 0 \qquad P(|\overline{X_n} - \mu| \geq \varepsilon)$$

$$= P(|\overline{X_n} - \mu|^2 \geq \varepsilon^2)$$

$$\leq \frac{1}{\varepsilon^2} E|\overline{X_n} - \mu|^2 \quad \leftarrow \text{Chebychev's inequality (Sec 3.5)}$$

$$= \frac{1}{\varepsilon^2} \operatorname{var} \overline{X_n}$$

$$= \frac{1}{\varepsilon^2} \frac{1}{n} \sigma^2 \to 0 \text{ as } n \to \infty.$$

$\square$

Def we say $X_n \to X$ almost surely (also, with probability 1)

if $$\boxed{P\left(\lim_{n\to\infty}|X_n - X| = 0\right) = 1.} \quad (*)$$

$$\Leftrightarrow \quad P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

Remark.

- Equivalently, $\forall \varepsilon > 0$ $P\left(\lim_{n\to\infty}|X_n - X| < \varepsilon\right) = 1.$

if this is true then

$$P\left(\lim_{n\to\infty}|X_n - X| < \frac{1}{2^k},\ \forall k \in \mathbb{N}\right)$$

$$= 1 - P\left(\exists k \in \mathbb{N},\ \lim_{n\to\infty}|X_n - X| \geq \frac{1}{2^k}\right)$$

$$\geq 1 - \sum_{k=0}^{\infty} P\left(\lim |X_n - X| \geq \frac{1}{2^k}\right)$$

$$= 1$$

So, $(*)$ is true.

- if $X_n \to X$ a.s. then $X_n \to X$ in probability.

if $X_n \to X$ a.s. then $\forall \varepsilon > 0$

$$1 = P\left(\lim_{n\to\infty} |X_n - X| < \varepsilon\right) = \mathbb{E}\,\mathbb{1}\left(\lim_{n\to\infty} |X_n - X| < \varepsilon\right)$$

$$\leq \lim_{n\to\infty} P\left(|X_n - X| < \varepsilon\right) \quad \text{(Fatou's lemma)}$$

Hence $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$.

Example

Useful to use the original def. of Random variable as a function on S: $s \mapsto X(s)$

Let $S = [0,1]$, P uniform dist. on S.

Define $\begin{cases} X_n(s) = s + s^n \\ X(s) = s \end{cases}$

Then $X_n(s) \longrightarrow X(s)$ $\forall s \in [0,1)$

So $P(s: \lim X_n(s) = X(s)) = P([0,1)) = 1$.

eventhough $P(X_n \neq X) = P(s \neq 0) = 1$.

Example

Let $X_1(s) = s + \mathbb{1}(s \in [0,1])$.

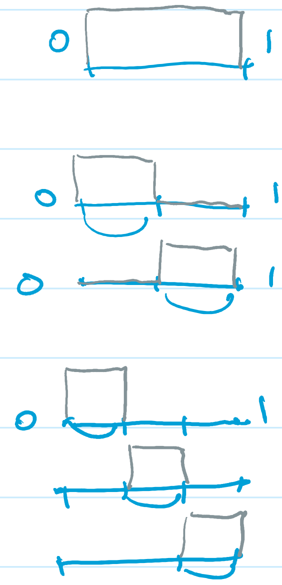
$X_2(s) = s + \mathbb{1}(s \in [0, \tfrac{1}{2}])$

$X_3(s) = s + \mathbb{1}(s \in [\tfrac{1}{2}, 1])$

$X_4(s) = s + \mathbb{1}(s \in [0, 1/3])$

$X_5(s) = s + \mathbb{1}(s \in [1/3, 2/3])$

$X_6(s) = s + \mathbb{1}(s \in [2/3, 1])$



Then

$$P(s : |X_n(s) - X(s)| > \varepsilon) \longrightarrow 0 \qquad \forall \varepsilon > 0$$

but $P(s : X_n(s) \to X(s)) = 0$.

Hence $X_n \to X$ in probability,

but $X_n \not\to X$ almost surely.

Example

Let $X = N(0,1)$ and $Y = -X$.

Then $X = Y$ in distribution,

but $P(X \neq Y) = 1$.

Remarks

Convergence a.s. $\Rightarrow$ Convergence in probability

Convergence in probability $\Rightarrow$ Convergence in distribution

Theorem Strong law of large numbers (SLLN)

if $X_1, X_2, \ldots$ are iid Random variables with $\mu = \mathbb{E}X_i$ and $\mathbb{E}|X_i| < \infty$ Then

$$\bar{X}_n = \frac{1}{n}(X_1 + \ldots + X_n) \longrightarrow \mu \quad \text{almost surely}$$

Remarks

- Proof for this is Similar in spirit, but technically more involved than that of the WLLN
- The condition $\mathbb{E}|X_i| < \infty$ is very mild.

Theorem Central Limit Theorem (CLT).

If $X_1, X_2, \ldots$ are iid Random variables whose MGFs exist in a neighborhood of 0.

Let $\mu = \mathbb{E} X_i$, $\sigma^2 = \operatorname{var} X_i > 0$ and $\bar{X}_n = \frac{1}{n}(X_1 + \ldots + X_n)$. Let $Z \sim N(0,1)$.

Then $\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \longrightarrow Z$ in distribution.

Remark

- Perhaps the most celebrated theorem in probability
- widely applicable : only $\operatorname{var} X_i < \infty$ is required
- Variations of CLT is possible !

Let $Y_i = (X_i - \mu)/\sigma$. Then $\mathbb{E} Y_i = 0$, $\mathbb{E} Y_i^2 = 1$.

So
$$\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) = \frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + \ldots + X_n - n\mu}{n}\right)$$
$$= \frac{1}{\sqrt{n}}(Y_1 + \ldots + Y_n).$$

We may write $\frac{1}{\sqrt{n}}(Y_1 + \ldots + Y_n) \xrightarrow{d} N(0,1)$.

Proof Sketch.

- The MGF for Z is $M_Z(t) = \mathbb{E}e^{tZ} = e^{\frac{1}{2}t^2}$.
  It is enough to show that

$$M_{\frac{1}{\sqrt{n}}(Y_1+\dots+Y_n)}(t) \longrightarrow M_Z(t) \quad \text{as } n \to \infty$$

$$\forall\ |t| < \delta \quad \text{for some } \delta > 0.$$

- Now

$$M_{\frac{1}{\sqrt{n}}(Y_1+\dots+Y_n)}(t) = \left(M_Y(t/\sqrt{n})\right)^n.$$

Apply Taylor expansion, which is valid for small $|t/(\sqrt{n}\sigma)|$.

$$M_Y(t/\sqrt{n}) = \underbrace{M_Y(0)}_{1} + \underbrace{M_Y^{(1)}(0)}_{\mathbb{E}Y=0}\frac{t}{\sqrt{n}} + \frac{1}{2}\underbrace{M_Y^{(2)}(0)}_{\mathbb{E}Y^2=1}\frac{t^2}{n} + o\left(\frac{t^2}{n}\right)$$

where $o\left(\frac{t^2}{n}\right)$ vanishes faster than $\frac{t^2}{n}$ as $n\to\infty$.

- Thus

$$\left(M_Y(t/\sqrt{n})\right)^n = \left(1 + \frac{1}{2}\frac{t^2}{n} + o\left(\frac{t^2}{n}\right)\right)^n$$

$$\longrightarrow \lim_{n\to\infty}\left(1 + \frac{1}{2}\frac{t^2}{n}\right)^{\frac{2n}{t^2}\cdot\frac{t^2}{2}} = e^{\frac{t^2}{2}}$$

$\square$