Limits for a sequence of random variables

$X_1, X_2, \dots$ i.i.d sequence of r.v.

independent, identically distributed

$S_n = X_1 + \dots + X_n$

Strong Law of Large Numbers:

if $\mu = EX_1 \in \mathbb{R}$ then

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu$$

Central Limit Theorem

if $\mu = EX_1 \in \mathbb{R}$, $\sigma^2 = \mathrm{Var}(X_1) < \infty$

then $\frac{S_n - n\mu}{\sqrt{n\sigma^2}}$ converges in distribution

to the standard normal distribution.

Named distributions

Discrete: Bernoulli, Binomial (n, p), Geometric (p)

negative binomial (r,p)

All of these can be represented in a prob space with an iid sequence of Bernoulli(p) random variables. $X_1, X_2, \ldots$ iid

$P(X_1 = 1) = p \quad P(X_1 = 0) = 1-p$

Poisson (λ) distribution $\lambda > 0$

range of possible values : $\{0, 1, 2, \ldots\}$

$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ $\quad \boxed{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1}$

Abs. continuous distributions

- Uniform $[a,b]$ $\quad a<b$

$$f(x) = \frac{1}{b-a} \mathbb{1}_{(x \in [a,b])}$$

- Exponential (λ) $\quad \lambda > 0$

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{(x>0)}$$

- Gamma(a, λ) $\quad a, \lambda > 0$

$$f(x) = \frac{\lambda^a x^{a-1}}{\Gamma(a)} e^{-\lambda x} \mathbb{1}(x>0)$$

- Normal distribution $\mu \in \mathbb{R}$, $\sigma > 0$
  $\mathcal{N}(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad x \in \mathbb{R}$$

Convergence of random variables

$X_1, X_2, \ldots, X$ random variables on the same probability space

1) We say that $X_n \xrightarrow{a.s.} X$ as $n \to \infty$ if $\underbrace{P(\lim_{n\to\infty} X_n = X)}$ $= 1$.

almost sure convergence / convergence with probability one

2) We say that $X_n \xrightarrow{P} X$ as $n \to \infty$ if for every $\varepsilon > 0$

$$\lim_{n\to\infty} P(\underbrace{|X_n - X| > \varepsilon}) = 0$$

convergence in probability / stochastic convergence

3) We say that $X_n \xrightarrow{L_p} X$ for $p>0$

if $E\left[|X_n - X|^p\right] \to 0$ as $n \to \infty$.

L_p convergence

If $X = c$ is a constant then we can make sense $X_n \xrightarrow{P} c$, $X_n \xrightarrow{L_p} c$ even in the case when $X_1, X_2, \dots$ do not live on the same prob space.

Lemma: 1) If $X_n \xrightarrow{a.s.} X$ then $X_n \xrightarrow{P} X$.

2) If for a $p>0$ $X_n \xrightarrow{L_p} X$ then $X_n \xrightarrow{P} X$.

3) The other implications do not hold in general.

Proof: 1) We need:

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

We know $X_n \xrightarrow{a.s.} X$.

$P(|X_n - X| \geq \varepsilon) = E\left[\mathbb{1}_{\{|X_n - X| \geq \varepsilon\}}\right] \longrightarrow 0$

$\mathbb{1}_{\{|X_n - X| \geq \varepsilon\}} \xrightarrow{a.s.} 0$, they are bounded

hence by the Dominated Conv. Theorem

we have $E\left[\mathbb{1}_{\{|X_n - X| \geq \varepsilon\}}\right] = E[0] = 0$

2) $E\left[|X_n - X|^p\right] \to 0$

We need to control $P(|X_n - X| \geq \varepsilon)$.

$$P(|X_n - X| \geq \varepsilon) = P(|X_n - X|^p \geq \varepsilon^p) \overset{\text{Markov}}{\leq}$$

$$\leq \frac{E\left[|X_n - X|^p\right]}{\varepsilon^p} \xrightarrow[n \to \infty]{} 0$$

Weak Law of Large Numbers

Thm: $X_1, X_2, \ldots$ iid random variables

with $EX_1 \in \mathbb{R}$, $EX_1^2 < \infty$.

Then $\dfrac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{P} E[X_1]$.

Proof: We will prove that $\dfrac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{L_2} E[X_1]$.

Notation: $\mu = E[X_1]$, $\sigma^2 = Var(X_1)$

Goal: $E\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - \mu\right|^2\right] \to 0$ as $n \to \infty$.

$$E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\left(E[X_1] + \dots + E[X_n]\right)$$

$$= \frac{1}{n} n \cdot \mu = \mu$$

$$E\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - \underset{\uparrow \atop E\left[\frac{\sum X_i}{n}\right]}{\mu}\right|^2\right] = \mathrm{Var}\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)$$

$$= \frac{1}{n^2} \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$\mathrm{Var}(X+Y) = E\left[\left(X+Y - E[X+Y]\right)^2\right]$$

$$= E\left[\left(X - E[X] + Y - E[Y]\right)^2\right]$$

$$= E\left[(X-E[X])^2\right] + 2\underline{E\left[(X-E[X])(Y-E[Y])\right]}$$

$$+ E\left[(Y-E[Y])^2\right]$$

$$= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j)$$

If $X_1, X_2, \dots, X_n$ are independent then $\mathrm{Cov}(X_i, X_j) = 0$ if $i \neq j$ and $\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$.

$$E\left[\left(\frac{\sum_{i=1}^{n} X_i}{n} - \mu\right)^2\right] = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$

$$\mathrm{Var}\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{\sigma^2}{n} \longrightarrow 0 \quad \text{as } n \to \infty$$

The statement holds for $X_1, X_2, \dots$ which are uncorrelated, $E(X_n) = \mu$, $\mathrm{Var}(X_n) = \sigma^2$.

Application: the Weierstrass approx thm

Thm: $f:[0,1]\to\mathbb{R}$, continuous. Then there are polynomials f_n such that $f_n \to f$ uniformly on $[0,1]$.

Proof: $$f_n(x) \overset{def}{=} \sum_{i=0}^{n} \underbrace{\binom{n}{i} x^i (1-x)^{n-i}} f\left(\frac{i}{n}\right)$$

This is a polynomial. We will show that
$$f_n \to f \text{ uniformly on } [0,1].$$

Fix $p \in [0,1]$, let $X_1, X_2, \dots, X_n, \dots$ be iid Bernoulli(p). Then $S_n = X_1 + \dots + X_n$ has Binomial(n,p).

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$
$$k = 0, \dots, n$$

$$E[X_1 + \dots + X_n] \overset{\text{linearity}}{=} E[X_1] + \dots + E[X_n] = np$$

$$\mathrm{Var}(X_1 + \dots + X_n) \underset{\text{independence}}{=} \mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_n)$$
$$= n\,\mathrm{Var}(X_1)$$
$$= n(p-p^2) = np(1-p)$$

$X_1 \sim$ Bernoulli(p)
$E[X_1] = p \quad E[X_1^2] = p$
$\mathrm{Var}(X_1) = E[X_1^2] - (E[X_1])^2 = p - p^2$

$$E\left[f\left(\frac{S_n}{n}\right)\right] = f_n(p)$$

$$E[f(X)] = \int f(x)\, dQ_X$$

$$f_n(x) = \sum_{i=0}^{n} \underbrace{\binom{n}{i} x^i (1-x)^{n-i}} f\left(\frac{i}{n}\right)$$

$$E\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{i=0}^{n} f\left(\frac{i}{n}\right) P(S_n = i)$$

$$\left|f_n(p) - f(p)\right| = \left|E\left[f\left(\frac{S_n}{n}\right)\right] - f(p)\right|$$

$$= \left|E\left[f\left(\frac{S_n}{n}\right) - f(p)\right]\right| \le E\left[\underbrace{\left|f\left(\frac{S_n}{n}\right) - f(p)\right|}\right] =$$

Fix $\delta > 0$

this has small probability

$$= E\left[\left|f\left(\frac{S_n}{n}\right) - f(p)\right| \mathbb{1}\left(\left|\frac{S_n}{n} - p\right| \ge \delta\right)\right]$$

$$+ E\left[\left|f\left(\frac{S_n}{n}\right) - f(p)\right| \mathbb{1}\left(\left|\frac{S_n}{n} - p\right| < \delta\right)\right]$$

this is small