$X_n \xrightarrow{a.s.} X$ if $P(\underbrace{\lim_{n\to\infty} X_n = X}_{\text{this depends on } X,\ X_n,\ n\geq 1}) = 1.$

Borel - Cantelli Lemmas

Def: $A_1, A_2, \ldots$ are events

$$\limsup_{n\to\infty} A_n = \lim_{m\to\infty} \underbrace{\bigcup_{n=m}^{\infty} A_n}_{\substack{\text{this is a}\\ \text{decreasing sequence}\\ \text{of events}}} = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n$$

$= \{\omega : \omega \in A_n \text{ for infinitely many } n\}$

$= \{A_n \text{ i.o.}\}$

$$\liminf_{n\to\infty} A_n = \lim_{m\to\infty} \underbrace{\bigcap_{n=m}^{\infty} A_n}_{\text{increasing seq}} = \bigcup_m \bigcap_{n=m} A_n$$

$= \{\omega : \omega \in A_n \text{ for all, but finitely many } n\}$

$$1_{\limsup\limits_{n\to\infty} A_n} = \limsup_{n\to\infty} 1_{A_n}, \quad 1_{\liminf\limits_{n\to\infty} A_n} = \liminf_{n\to\infty} 1_{A_n}$$

(First)

· Borel–Cantelli lemma:

$A_1, A_2, \ldots$ are events with $\sum_{n=1}^{\infty} P(A_n) < \infty$.

Then $P(A_n \text{ i.o.}) = 0$.

Proof: 1) Set $N = \sum_{n=1}^{\infty} 1_{A_n} \in \{0,1,\ldots\} \cup \{\infty\}$

$\{N = \infty\} = \{A_n \text{ i.o.}\}$

$$E[N] \underset{\uparrow}{=} \sum_{n=1}^{\infty} \underset{\uparrow}{P(A_n)} < \infty$$

$E\left[\sum_{n=1}^{\infty} 1_{A_n}\right]$ $\quad \sum_{n=1}^{\infty} E[1_{A_n}]$

If $E[N] < \infty$ then $\underset{P(A_n \text{ i.o.})}{\underset{\shortparallel}{P(N=\infty)}} = 0$

2) $\{A_n \text{ i.o.}\} \subseteq \bigcup_{n=m} A_n$

$$P(A_n \text{ i.o.}) \leq P\left(\bigcup_{n=m} A_n\right) \leq \sum_{n=m}^{\infty} P(A_n)$$

$\sum_{n=1}^{\infty} P(A_n) < \infty$ so $\lim_{m\to\infty} \sum_{n=m}^{\infty} P(A_n) = 0$

Hence $P(A_n \text{ i.o.}) = 0$. ∎

Lemma: $X_1, X_2, \dots$ and X are random variables. $\varepsilon_n \to 0$ as $n \to \infty$

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

implies $X_n \xrightarrow{a.s.} X$.

Proof: $A_n = \{|X_n - X| > \varepsilon_n\}$

$\sum_{n=1}^{\infty} P(A_n) < \infty$, hence by BC-1. we have $P(A_n \text{ i.o.}) = 0$.

$\{A_n \text{ i.o.}\}^c = \{\omega : \omega \in A_n^c \text{ for } n \text{ large enough}\}$

$= \{\omega : |X_n(\omega) - X(\omega)| \leq \varepsilon_n \text{ for } n \text{ large enough}\}$

on this event $X_n \to X$.

$P(A_n^c \text{ holds for } n \text{ large enough}) = 1$

$\rightsquigarrow \quad P(X_n \to X) = 1$

Strong Law of Large Numbers with 4th moment

$X_1, X_2, \dots$ i.i.d $\quad \boxed{E[X_1^4] < \infty}$

we will remove this later

$S_n = X_1 + \dots + X_n$

Then $\frac{S_n}{n} \xrightarrow{a.s.} E[X_1]$.

Proof: $\mu = E[X_1] \qquad E\left[\frac{S_n}{n}\right] = \mu$.

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon_n\right) \leq \frac{E\left[\left|\frac{S_n}{n} - \mu\right|^4\right]}{\varepsilon_n^4}$$

$$E\left[\left(\frac{S_n}{n} - \mu\right)^4\right] = E\left[\frac{1}{n^4}(S_n - n\mu)^4\right]$$

$$= \frac{1}{n^4} E\left[\left(\sum_{j=1}^n (X_j - \mu)\right)^4\right]$$

$$E\left[\left(\sum_{j=1}^n (X_j - \mu)\right)^4\right] =$$

$$= E\Bigg[\underbrace{\sum_{j=1}^n (X_j - \mu)^4}_{1.} + c_1 \underbrace{\sum_{i,j} (X_i - \mu)(X_j - \mu)^3}_{2.}$$

$$+ c_2 \underbrace{\sum_{i,j} (X_i - \mu)^2 (X_j - \mu)^2}_{3.} + c_3 \underbrace{\sum_{i,j,k} (X_i - \mu)(X_j - \mu)(X_k - \mu)^2}_{4.}$$

$$+ c_4 \underbrace{\sum_{i,j,k,\ell} (X_i - \mu)(X_j - \mu)(X_k - \mu)(X_\ell - \mu)}_{5.}\Bigg]$$

$$= E\left[\sum_{j=1}^{n}(X_j-\mu)^4 + c_2 \sum_{i,j}(X_i-\mu)^2(X_j-\mu)^2\right]$$

$$= n\,E\left[(X_1-\mu)^4\right] + c_2\binom{n}{2}\underbrace{E\left[(X_1-\mu)^2\right]E\left[(X_1-\mu)^2\right]}_{(\mathrm{Var}\,X_1)^2}$$

$$\le C\cdot n^2 \quad \text{for some } C>0.$$

$$E\left[\left(\frac{S_n}{n}-\mu\right)^4\right] = \frac{1}{n^4}E\left[(S_n-n\mu)^4\right]$$

$$\le \frac{C}{n^2}$$

$$P\left(\left|\frac{S_n}{n}-\mu\right| > \varepsilon_n\right) \le \frac{C}{n^2\,\varepsilon_n^4}$$

With $\varepsilon_n = n^{-\frac{1}{8}}$ then $\sum P\left(\left|\frac{S_n}{n}-\mu\right|>\varepsilon_n\right)<\infty$

and $\varepsilon_n \to 0$ hence

$$\frac{S_n}{n}-\mu \xrightarrow{a.s.} 0.$$

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

Application: $X \sim \text{Uniform}[0,1]$

$X = 0.x_1x_2\ldots\ldots$ $\boxed{x_1, x_2, \ldots \text{ iid with distribution that is uniform on } \{0,1,\ldots,9\}}$

Then $\dfrac{\sum_{i=1}^{n} x_i}{n} \xrightarrow{a.s.} \dfrac{9}{2}$.

Moreover $\dfrac{\sum_{i=1}^{n} \mathbb{1}\{x_i = a\}}{n} \xrightarrow{a.s.} \dfrac{1}{10}$

for any $a \in \{0,1,\ldots,9\}$

Ex: $X_1, X_2, \ldots$ iid Bernoulli(p)

$$Y = \sum_{k=1}^{\infty} X_k \frac{1}{2^k} \qquad p \neq \frac{1}{2}$$

The distribution of Y is singular with respect to Lebesgue.

$$P\left(\frac{\sum_{i=1}^{n} X_i}{n} \to \frac{1}{2}\right) = 0, \text{ but}$$

$\uparrow p \neq \frac{1}{2}$

$\dfrac{\sum X_i}{n} \to \dfrac{1}{2}$ if $Y \sim \text{Unif}[0,1]$

We have seen that if $X_n \xrightarrow{a.s.} X$ then $X_n \xrightarrow{P} X$.

Theorem: $X_n \xrightarrow{P} X$ if and only if for any subsequence of n there is a further subsequence along which $X_{n_{a_k}} \xrightarrow[k\to\infty]{a.s.} X$.

Proof: Assume that the first subsequence is the whole sequence.

First choose $\varepsilon_k \to 0$. Then we can find n_k so that $P(|X_{n_k}-X|>\varepsilon_k) \le \frac{1}{2^k}$, with $n_1 < n_2 < n_3 < \dots$.

Then $\sum_{k=1}^{\infty} P(|X_{n_k}-X|>\varepsilon_k) < \infty$, hence by BC-1 we have $X_{n_k} \xrightarrow{a.s.} X$.

For the other direction: assume that $X_n \overset{P}{\nrightarrow} X$. Then there is $\varepsilon>0$ so that $\lim P(|X_n-X|>\varepsilon) \neq 0$ so there is a subsequence with

$$\liminf_{k \to \infty} P(|X_{n_k} - X| > \varepsilon) \geq 0$$

This subsequence will not have an almost surely convergent subsequence.

Lemma: $y_n, n \geq 1$ is a sequence in a topological space. If every subsequence has a further subsequence along which we converge to a value y then $y_n \to y$.

This means that a.s. convergence of random variables cannot come from a metric!

Convergence in probability can be described with the metric

$$\delta(X, Y) = E\left[\frac{|X-Y|}{1+|X-Y|}\right].$$

If $\delta(X_n, X) \to 0$ then $X_n \xrightarrow{P} X$.

Lemma: If f is continuous, $X_n \xrightarrow{P} X$ then $f(X_n) \xrightarrow{P} f(X)$. If f is bounded then $E[f(X_n)] \to E[f(X)]$.