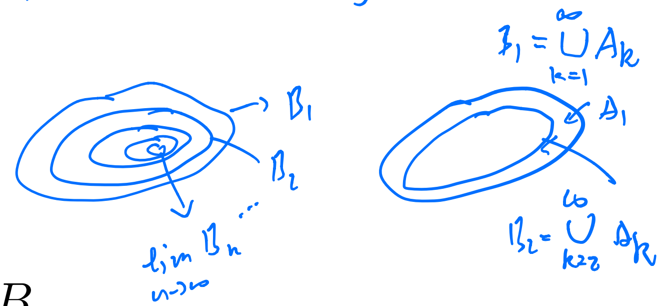
$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$

$A_n = \{\omega_2, \omega_3, \omega_5\}$ $\quad \mathcal{F} = \{\{\omega_1, \omega_2\} \cdots, \{\omega_2, \omega_3, \omega_5\} \cdots\}$



Let $\Omega \supset A_n \in \mathcal{F}$.

Define $B_n = \cup_{k \geq n} A_k$ and

$n \uparrow, \ B_n \downarrow$

$$\text{limsup} A_n = \bigcap_{n=1}^{\infty} B_n.$$

"limsupA_n" represents the event that infinitely many A_n take place. Sometimes we write $\{A_n\ ,\ i.o.\}$ to denote that.

limsup A_n happens, then infinitely many A_n happens

Note that $B_n \downarrow \text{limsup} A_n$ and therefore by laws of probability,

$$P(\text{limsup} A_n) = lim_{n\to\infty} P(B_n).$$

Theorem 1.1.1 (iii) Prob Durrett.

$S = \sum_{k=1}^{\infty} P(A_k) < \infty \qquad S_n = \sum_{k=1}^{n} P(A_k)$

No independence assumed.

$\sum_{k=n}^{\infty} P(A_k) = S - S_{n-1}$

Fact: 1st Borel-Cantelli lemma. $\limsup_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = \limsup_{n\to\infty} S - \limsup_{n\to\infty} S_{n-1} = 0$

If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\text{limsup} A_n) = 0.$

$$Proof.\ 0 \le P(\text{limsup} A_n) = lim_{n\to\infty} P(B_n)$$

$$= lim_{n\to\infty} P(\cup_{k\ge n} A_k) \le \overline{lim}_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0.$$

$$\le \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) \quad \le \limsup_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$$

① 放缩两次，

② 取 lim sup，避免不收敛的情况。

③ limsup ⇒ >, >

Almost sure convergence.

Recall: $X_n \to X$, *almost surely* iff
$P\{\omega$: For every $\epsilon > 0$, there exists $N(\omega, \epsilon)$, such that
$|X_n - X| \le \epsilon$, for all $n \ge N(\omega, \epsilon)\} = 1$

$$P(\lim_{n\to\infty} X_n = X) = 1$$

Equivalent

Claim: It is enough to establish the following seemingly weaker statement: For each ϵ>0, $P\{\omega :$ There exists $N(\omega, \epsilon)$, such that $|X_n - X| \leq \epsilon$, for all $n \geq N(\omega, \epsilon)\} = 1.$

Why?

Because we can interest countably infinitely many sure events to get a sure event. (Prove at home)

Let $k \geq 1$ be an arbitrary integer. Use $\epsilon = 1/k.$ Then we get $P\{\omega :$ There exists $N(\omega, k)$, such that $|X_n - X| \leq 1/k,$ for all $n \geq N(\omega, k)\} = 1.$

Hence $P\,(\cap_k\{\omega :$ There exists $N(\omega, k)$, such that $|X_n - X| \leq 1/k,$ for all $n \geq N(\omega, k)\}) = 1$

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - P\left[\left(\bigcap_{n=1}^{\infty} A_n\right)^c\right] = 1 - P\left[\bigcup_{n=1}^{\infty} (A_n)^c\right] \geq 1 - \sum_{n=1}^{\infty} P[A_n^c]$$

$$P(A_n) = 1 \qquad = 1 - \sum_{n=1}^{\infty} 0$$

$$= 1$$

Note $\cap_k \{\omega :$ There exists $N(\omega, k)$, such that $|X_n - X| \leq 1/k$, for all $n \geq N(\omega, k)\} = \cap_{\epsilon>0} \{\omega :$ There exists $N(\omega, \epsilon)$, such that $|X_n - X| \leq \epsilon$, for all $n \geq N(\omega, \epsilon)\}$

$P(\cdot)=1$

Hence we have established the seemingly stronger statement in the definition of almost sure convergence.

Suppose $X_n \xrightarrow{P} X$, which means for each $\epsilon > 0$, $P(|X_n - X| > \epsilon) \to 0$. However it is hard to combine the sets $\{|X_n - X| > \epsilon\}$ without knowing how fast these probabilities go to zero.

$P(|X_n - X| > \varepsilon, i.o.) = 0$

$P(\lim_{n\to\infty} X_n(\omega) = X(\omega)) = 1$ 逐点收敛. 不相等的点测度为0
$(\mathbb{R}, \mathcal{B})$

$\forall \varepsilon > 0$, $\lim_{n\to\infty} P(|X_n(\omega) - X(\omega)| < \varepsilon) = 1$ 依测度收敛, 不等的点$\to 0$, 但不一定测度为0. 可以是$\to 0$的有测度的点.
$(\mathbb{R}, \mathcal{B}\to 0)$

Complete Convergence

Suppose for each $\epsilon > 0$, $\sum_{n=1}^{\infty} P(|X_n - X| > \epsilon) < \infty$. This is called complete convergence; we write $X_n \xrightarrow{c} X$.

$P(\lim_{n\to\infty} X_n \neq X, i.o.) = 0$

Note that by, Borel-Cantelli lemma, for each $\epsilon > 0$, $P(|X_n - X| > \epsilon,\ i.o.) = 0$ and hence $X_n \xrightarrow{\text{a.s.}} X$.

故 $P(\cdot) = 0$，说明只有有限个 $|X_n - X| > \varepsilon$ happen. 故 $P(\lim_{n\to\infty} X_n = X) = 1$

$E_n = |X_n - X| > \varepsilon$ E_n, i.o. happens, infinitely many $|X_n - X| > \varepsilon$ happen,

<u>We can show a.s. convergence implies convergence in probability.</u>

Since $P(|X_n - X| > \epsilon,\ i.o.) = 0$, we get $lim_{n\to\infty} P(B_n) = 0$ where $B_n = \cup_{k \geq n} A_k \supset A_n$, with $A_n = \{|X_n - X| > \epsilon\}$. Hence $P(A_n) \to 0$, showing $X_n \xrightarrow{P} X$.

$A_n = \{ |X_n - X| > \varepsilon \}$, $B_n = \bigcup_{k=n}^{\infty} A_k \supset A_n$

$X_n \xrightarrow{a.s.} X \Rightarrow P(|X_n - X| > \varepsilon, i.o.) = 0 \Rightarrow P(A_n, i.o.) = 0 \Rightarrow$

$P(\limsup_{n\to\infty} A_n) = 0 \Rightarrow \lim_{n\to\infty} P(B_n) = 0 \Rightarrow \lim_{n\to\infty} P(A_n) = 0 \Rightarrow \lim_{n\to\infty} (|X_n - X| > \varepsilon) = 0$

$P(X_n \le x) \to P(X \le x)$ for $\forall x$ continuous

(If the distribution is Normal, then all points are required to converge.

Convergence in distribution

X_n converges in distribution to X, written as $X_n \xrightarrow{d} X$, if

$$F_{X_n}(x) \to F_X(x), \text{ whenever } F_X \text{ is continuous at } x.$$

Why not at all x?
Because, that will be too much to ask as shown by the following example.

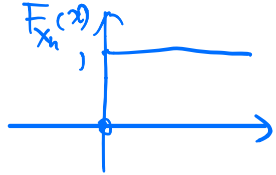
$$X_n = \frac{1}{n}, \text{ with probability } 1$$

$$X = 0, \text{ with probability } 1.$$

Then X_n converges to X even almost surely.

$P(\frac{1}{n} \le x) \overset{?}{\to} P(0 \le x)$ for $\forall x$

let $x=0$, $\underbrace{P(\frac{1}{n} \le 0)}_{=0}$ $\not\to$ $P(0 \le 0) = 1$



$n \to \infty$

$$F_{X_n}(0) = 0, \text{ for all } n \qquad \to 0$$

but $\quad F_X(0) = 1.$

Note that the limiting distribution function F_X is discontinuous at 0.

Normal continuous, every point

In general, X_n and X do not have to be defined on the same probability space.

不可以作差 $X_n - X$,

The most important example of distribution convergence in statistics is due to the central limit theorem (CLT).

$$Y_i,\ i \geq 1, \text{ i.i.d. } E(Y_i) = \mu,\ Var(Y_i) = \sigma^2.$$

Then

$$X_n = \frac{\overline{Y}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1) = X$$

Fact: Convergence in probability implies convergence in distribution.

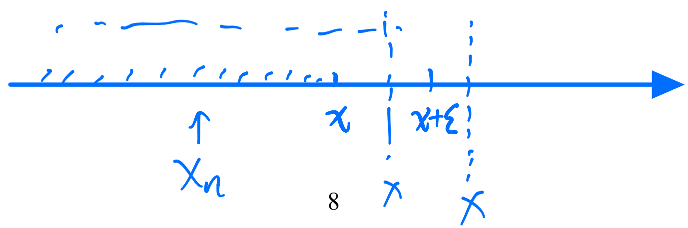
$$\begin{cases} X_n - X > \epsilon \\ X_n - X < -\epsilon \end{cases} \Rightarrow \begin{matrix} X < X_n - \epsilon \\ X > X_n + \epsilon \end{matrix}$$

Suppose F_X is continuous at x.

Note $\{X_n \leq x\} \subset \{X \leq x + \epsilon\} \cup \{|X_n - X| > \epsilon\}$

One of two events must happen.

Hence



$$\{X \leq x\} \subset \{X_n \leq x + \varepsilon\} \cup \{|X_n - X| > \varepsilon\}$$

$$F_X(x) \leq F_{X_n}(x+\varepsilon) + P(|X_n - X| > \varepsilon)$$

$$F_{X_n}(x) \leq F_X(x+\epsilon) + P(|X_n - X| > \epsilon)$$

$$limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x+\epsilon) + 0$$

Now let $\epsilon \downarrow 0$, and use the fact that F_X is continuous at x, to obtain

$$limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x).$$

Similarly, one can show (try this at home) that

$$liminf_{n\to\infty} F_{X_n}(x) \geq F_X(x)$$

Thus

$$F_X(x) \leq \liminf F_{X_n}(x+\varepsilon) + P(|X_n - X| > \varepsilon)$$

$$F_X(x) \leq \liminf F_{X_n}(x)$$

$$F_X(x) \leq liminf_{n\to\infty} F_{X_n}(x) \leq limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x)$$

– sandwich trick!

Hence

$$lim\, F_{X_n}(x) = F_X(x).$$

Distribution convergence does not imply convergence in probability.

Easy counter example: Suppose Z is a given random variable that is standard normally distributed.

Let $X_n = -Z$ and $X = Z$

Then

$$F_{X_n}(x) = \Phi(x), \text{ for all } n, \text{ and } x$$

$$F_X(x) = \Phi(x), \text{ for all } x$$

Hence

$$F_{X_n}(x) \longrightarrow F_X(x), \text{ for all } x$$

and $X_n \xrightarrow{d} X$.

However, $X_n - X = 2Z \stackrel{p}{\not\to} 0$.

Exc. If $X_n \xrightarrow{d} c$, where c is a constant, then $X_n \xrightarrow{p} c$.

$$P(X_n \le x) \xrightarrow{n\to\infty} P(c \le x)$$

To prove $\lim_{n\to\infty} P(|X_n - c| \le \varepsilon) = 1$

$$P(c-\varepsilon \le X_n \le c+\varepsilon)$$

$$= F_{X_n}(c+\varepsilon) - F_{X_n}(c-\varepsilon)$$

$$= P(X_n \le c+\varepsilon) - P(X_n \le c-\varepsilon)$$

$$= P(c \le c+\varepsilon) - P(c \le c-\varepsilon)$$

$$= P(0 \le \varepsilon) - P(0 \le -\varepsilon)$$

$$= 1 - 0 = 1$$