Comps Review - Probability Theory

Tianchen Qian 4/11/2013

Four kinds of convergence:

[(0) $X_n \to X$ sample point-wise (most naive)]

(1) $X_n \xrightarrow{a.s.} X$ (Strong Law of Large Numbers)

(2) $X_n \xrightarrow{P} X$ (Weak Law of Large Numbers)

(3) $X_n \xrightarrow{L^p} X$

(4) $X_n \xrightarrow{\mathcal{D}} X$ (Central Limit Theorem)

DEFINITIONS:

- $X_n \xrightarrow{a.s.} X$, if $P\{\omega: X_n(\omega) \not\to X(\omega)\} = 0$;
  iff $\forall \varepsilon > 0$, $P\{|X_n - X| > \varepsilon \text{ i.o.}\} = 0$. (Th 4.2.2)

- $X_n \xrightarrow{P} X$, if $\forall \varepsilon > 0$, $P\{|X_n - X| < \varepsilon\} \longrightarrow 1$, as $n \to \infty$
  Alternative notation: $X_n - X = o_p(1)$.

- $X_n \xrightarrow{L^p} X$. if $\|X_n - X\|_p \to 0$, as $n \to \infty$,
  i.e. $E|X_n - X|^p \to 0$, as $n \to \infty$.
  - DEF of L^p-norm: $\|X\|_p := (E|X|^p)^{\frac{1}{p}}$, for $0 < p < \infty$.

- $X_n \xrightarrow{\mathcal{D}} X$, if $\underbrace{F_n \Rightarrow F}_{\text{corresponding distribution function}}$, i.e. $F_n(x) \to F(x)$ at all continuous points of F (i.e. when F is cont. at x),
  or $X_n \xrightarrow{\mathcal{L}} X$, $X_n \xrightarrow{d} X$
  iff $\underbrace{\mu_n \Rightarrow \mu}_{\text{corresponding distribution}}$, i.e. $\exists$ dense set $D \subset \mathbb{R}^1$, such that $\forall a < b \in D$, $\mu_n(a,b] \to \mu(a,b]$
  - $\mu_n \Rightarrow \mu$ iff $\forall a, b$ such that $\mu\{a\} = \mu\{b\} = 0$,
    $\mu_n(a,b] \longrightarrow \mu(a,b]$. (Th 4.3.1)

RELATIONSHIP BETWEEN THEM:

$\boxed{\text{Figure 1.}}$ True implications.

$$\begin{tikzcd}
\text{a.s.} \arrow[dr, Rightarrow, "\text{Th 4.1.2}"'] & & L^p \arrow[dl, Rightarrow, "\text{Th 4.1.4}"] \\
& P \arrow[d, Rightarrow, "\text{Th. 4.4.5}"] & \\
& \mathcal{D} &
\end{tikzcd}$$

$\boxed{\text{Figure 2.}}$ (partial converse - converse holds with further assumptions)

$$\begin{tikzcd}
\text{a.s.} & & L^P \\
& P \arrow[ul, Rightarrow, "\substack{\text{along a subsequence} \\ \text{Th 4.2.3}}"] \arrow[ur, Rightarrow, "\substack{L^P\text{- domination} \\ \text{Th 4.1.4}}"'] & \\
& \mathcal{D} \arrow[u, Rightarrow, "\substack{\text{when } X \text{ is a constant} \\ \text{Ex 4.4.4}}"'] &
\end{tikzcd}$$

<u>THM 4.2.3</u>: If $X_n \xrightarrow{P} X$, then $\exists\, (n_k)$ such that $X_{n_k} \xrightarrow{a.s.} X$ (as $k \to \infty$).

<u>THM 4.1.4</u>: If $X_n \xrightarrow{P} X$ and $\{X_n\}$ is L^P-dominated, then $X_n \xrightarrow{L^P} X$.
(only true for $0<p<\infty$)

- Distinguish between L^P-dominated and L^P-bounded:
  - $\{X_n\}$ is L^P-dominated, if $\exists$ random variable $Y \in L^P$, such that $|X_n| \leq Y$, $\forall n$.
  - $\{X_n\}$ is L^P-bounded, if $\exists$ (large but finite) number $M \in \mathbb{R}$, such that $\|X_n\|_p \left(= [E(X_n)^p]^{1/p}\right) \leq M$, $\forall n$.
    (equivalently, if $\sup_n \|X_n\|_p < \infty$).

  L^P-domination $\Rightarrow$ L^P-boundedness. But the converse is false.

<u>EX 4.4.4</u>: If $X_n \xrightarrow{\mathcal{D}} a \in \mathbb{R}$, then $X_n \xrightarrow{P} a$.

Counter examples to Fig 2 without further assumptions.

Example 1 ("circus") $X_n \xrightarrow{L^P} X$, but $X_n \not\xrightarrow{a.s.} X$;
$X_n \xrightarrow{P} X$, but $X_n \not\xrightarrow{a.s.} X$.

$\Omega = (0,1]$, $\mathcal{F} = \mathcal{B}$ = Borel on $(0,1]$, P = Lebesgue measure on $(0,1]$.

$X_0 = I_{(0,1]}$,

$X_1 = I_{(0,\frac{1}{2}]}$, $X_2 = I_{(\frac{1}{2},1]}$,

$X_3 = I_{(0,\frac{1}{3}]}$, $X_4 = I_{(\frac{1}{3},\frac{2}{3}]}$, $X_5 = I_{(\frac{2}{3},1]}$,

$X_6 = I_{(0,\frac{1}{4}]}$, $X_7 = I_{(\frac{1}{4},\frac{2}{4}]}$, $X_8 = I_{(\frac{2}{4},\frac{3}{4}]}$, $X_9 = I_{(\frac{3}{4},1]}$,

...

$X_n \xrightarrow{L^P} 0$, $X_n \xrightarrow{P} 0$,

But for each sample point, $X_n(\omega)$ has no limit.

Example 2 ("skyscrapers") $X_n \xrightarrow{a.s.} X$, but $X_n \not\xrightarrow{L^P} X$;
$X_n \xrightarrow{P} X$, but $X_n \not\xrightarrow{L^P} X$.

$\Omega = (0,1]$, $\mathcal{F} = \mathcal{B}$, P = Lebesgue measure on $(0,1]$.

$X_n = 2^n I_{(0,\frac{1}{n}]}$.

Then $X_n \to 0$ pointwise

$\Rightarrow X_n \xrightarrow{a.s.} 0$

$\Rightarrow X_n \xrightarrow{P} 0$.

But $\|X_n\|_p^p = \frac{2^{np}}{n} \to \infty$, so $\|X_n\|_p \to \infty$, so $\{X_n\}$ has no L^P- limit.

Example 3 $X_n \xrightarrow{\mathcal{D}} X$, but $X_n \not\xrightarrow{P} X$

$X_n, \mu_n := \sum_{k=1}^{n} \frac{1}{n} \delta_{\frac{k}{n}}$, X, μ : uniform $(0,1]$.

$\mu_n \Rightarrow \mu$, i.e. $X_n \xrightarrow{\mathcal{D}} X$,

but $P\{|X - X_n| > \frac{1}{2}\} = 1$, $\forall n$. so $X_n \not\xrightarrow{P} X$.

($\Omega = (0,1]$, $\mathcal{F} = \mathcal{B}$, P = Lebesgue measure on $(0,1]$).

IMPORTANT / USEFUL RESULTS

THM (Jensen's inequality) If φ is a convex function, then $\varphi(EX) \leq E[\varphi(X)]$.

THM (Markov's inequality) $\forall a > 0$,

$$P(|X| \geq a) \leq \frac{E|X|}{a}.$$

(useful in controlling tail probability by moments)

- Cor: $P(|X| \geq a) = P(|X|^p \geq a^p) \leq \frac{E|X|^p}{a^p}$.

THM (Slutsky's theorem)

If $X_n \xrightarrow{\mathcal{D}} X$, $Y_n \xrightarrow{\mathcal{D}} c \in \mathbb{R}$, then

- $X_n + Y_n \xrightarrow{\mathcal{D}} X + c$;
- $X_n Y_n \xrightarrow{\mathcal{D}} cX$;
- $Y_n^{-1} X_n \xrightarrow{\mathcal{D}} c^{-1} X$, if $c \neq 0$.

THM 4.5.4 If $0 < r < \infty$, $\forall n: X_n \in L^r$, and $X_n \xrightarrow{P} X$, then the following are equivalent:

(i) $(|X_n|^r)$ is uniformly integrable;

(ii) $X_n \xrightarrow{L^r} X$

(iii) $\|X_n\|_r \longrightarrow \|X\|_r < \infty$.

(relates $\xrightarrow{P}$, $\xrightarrow{L^p}$, convergence in L^p-norm and u.i.).