Stat 205A **September, 12, 2002**

Convergence of random variables, and the Borel-Cantelli lemmas

Lecturer: James W. Pitman *Scribes: Jin Kim (`jin@eecs`)*

1 Convergence of random variables

Recall that, given a sequence of random variables X_n, almost sure (a.s.) convergence, convergence in P, and convergence in L^p space are true concepts in a sense that $X_n \to X$. In this lecture, we will define weak convergence, or convergence in distribution, $\mathrm{P}_{X_n} \to \mathrm{P}_X$, which we write, by abuse of notation, $X_n \xrightarrow{d} X$.

Definition 1.1 (Convergence in distribution) *We say $X_n \xrightarrow{d} X$ if $\mathrm{P}(X_n \leq x) \longrightarrow \mathrm{P}(X \leq x)$ for all x at which the RHS is continuous.*

This weak convergence appears in the central limit theorem.

Theorem 1.2 $X_n \xrightarrow{d} X \iff \mathrm{E}f(X_n) \longrightarrow \mathrm{E}f(X)$ *for all bounded and continuous function f.*

Proof See Durrett. □

Theorem 1.3 *The following property holds among the types of convergence.*

$$\begin{tikzcd}
\boxed{X_n \xrightarrow{a.s.} X} \arrow[dr, Rightarrow] & & \boxed{X_n \xrightarrow{L^p} X} \arrow[dl, Rightarrow, "(*)"] \\
& \boxed{X_n \xrightarrow{p} X} \arrow[d, Rightarrow, "(**)"] & \\
& \boxed{X_n \xrightarrow{d} X} &
\end{tikzcd}$$

Proof $(*)$ can be proven by Chebychev inequality (with usually $p = 2$):

$$\mathrm{P}(\mid X_n - X \mid > \epsilon) \leq \frac{\mathrm{E} \mid X_n - X \mid^p}{\epsilon^p} \ ,$$

and $(**)$ is proven in Durrett. $\square$

Exercise. counter examples

- $\xrightarrow{p}$ but not $\xrightarrow{\text{a.s.}}$: moving blip
- $\xrightarrow{p}$ but not $\xrightarrow{L^p}$: try $X_n = n\mathbf{1}(0, 1/n)$. $X_n \xrightarrow{p} 0$ but $\mathrm{E} \mid X_n - 0 \mid = 1$, thus $X \not\longrightarrow 0$ in L^1.

Proposition 1.4 (Inducing a Metric) *$X_n \xrightarrow{a.s.} X$ cannot be metrized, but $X_n \xrightarrow{L^p} X$ and $X_n \xrightarrow{p} X$ can be metrized, e.g. using $\mathrm{E}(\mid X_n - X \mid \wedge 1)$. Furthermore, when so metrized, the space of random variables are complete.*

Proof See text (uses BCL).

Definition 1.5 (Infinitely Often (i.o.) and Eventually (ev.)) *Let q_n be some statement, e.g., $\mid X_n - X \mid > \epsilon$. We say ($q_n$ i.o.) if for all n, $\exists m \geq n \ : \ q_m$ is true, and (q_n ev.) if $\exists\, n \quad : \quad$ for all $m \geq n \ : \ q_m$ is true.*

Exercise. Note that the following holds;

- $X_n \longrightarrow X \quad \Longleftrightarrow \quad \forall \epsilon > 0, \mid X_n - X \mid < \epsilon$ ev.
- $X_n \not\longrightarrow X \quad \Longleftrightarrow \quad \forall \epsilon > 0, \mid X_n - X \mid > \epsilon$ i.o.
- $(q_n \text{ i.o.})^{\sim} \ = \ (q_n \text{ ev.})$

Similarly, for a sequence of events A_n in a prob space $(\Omega, \mathcal{F}, \mathrm{P})$, we can say the following;

- $(A_n \text{ i.o.}) = \{\omega : \omega \in A_n \ i.o.\} = \cap_n \cup_{m \geq n} A_m$
- $(A_n \text{ ev.}) = \{\omega : \omega \in A_n \ ev.\} = \cup_n \cap_{m \geq n} A_m$
- $(A_n \text{ i.o.})^c = A_n^c$ ev.

Main application of the idea of i.o. and ev. is to the proof of a.s. convergence. For example, since

$$(X_n \longrightarrow X) \quad = \quad \cap_{\epsilon > 0}(\mid X_n - X \mid < \epsilon \text{ ev.}) \ ,$$

we have

$$\mathrm{P}(X_n \longrightarrow X) \quad = \quad \lim_{\epsilon \longrightarrow 0} \mathrm{P}(\mid X_n - X \mid < \epsilon \text{ ev.}) \ .$$

Since the basic criterion for a.s. convergence can be written as

$$(X_n \longrightarrow X) \quad \Longleftrightarrow \quad \forall \epsilon > 0, \mathrm{P}(\mid X_n - X \mid > \epsilon \text{ i.o.}) = 0 \ ,$$

we are interested in conditions in some sequence of events A_n so that $\mathrm{P}(A_n)$ i.o. $= 0$.

2 Borel-Cantelli Lemma

Theorem 2.1 (Borel-Cantelli Lemma)

1. *If* $\sum_n \mathrm{P}(A_n) < \infty$*, then* $\mathrm{P}(A_n \text{ i.o.}) = 0$.

2. *If* $\sum_n \mathrm{P}(A_n) = \infty$ *and* A_n *are independent, then* $\mathrm{P}(A_n \text{ i.o.}) = 1$.

There are many possible substitutes for independence in BCL II, including Kochen-Stone Lemma.

Before prooving BCL, notice that

- $\mathbf{1}(A_n \text{ i.o.}) = \lim_{n\to\infty} \sup \mathbf{1}(A_n)$
- $\mathbf{1}(A_n \text{ ev.}) = \lim_{n\to\infty} \inf \mathbf{1}(A_n)$
- $(A_n \text{ i.o.}) = \lim_{m\to\infty} \mathrm{P}(\cup_{n>m} A_n)$ (as $m \uparrow$, $\cup_{n\geq m} A_n \downarrow$)
- $(A_n \text{ i.o.}) = \lim_{m\to\infty} \mathrm{P}(\cap_{n>m} A_n)$ (as $m \uparrow$, $\cap_{n\geq m} A_n \downarrow$).

Therefore,

$$\begin{aligned} \mathrm{P}(A_n \text{ ev.}) &\leq \lim \inf_{n\to\infty} \mathrm{P}(A_n) \qquad \text{by Fatou's lemma} \\ &\leq \lim \sup_{n\to\infty} \mathrm{P}(A_n) \qquad \text{obvious from definition} \\ &\leq \mathrm{P}(A_n \text{ i.o.}) \qquad \text{duel of Fatou's lemma (i.e. apply to } (\cdots)^\sim) \end{aligned}$$

Pf of BCL I

$$\begin{aligned} \mathrm{P}(A_n \text{ i.o.}) &= \lim_{m\to\infty} \mathrm{P}(\cup_{n\geq m} A_n) \\ &\leq \lim_{m\to\infty} \sum_{n\geq m}^{\infty} \mathrm{P}(A_n) \; = \; 0 \quad \text{since} \sum_{i=1}^{\infty} \mathrm{P}(A_n) < \infty. \; \square \end{aligned}$$

Pf of BCL I (Alternative method)
Consider a random variable $N := \sum \mathbf{1}(A_n)$, i.e. the number of events that occur. Then $\mathrm{E}[N] = \sum_{n=1}^{\infty} \mathrm{P}(A_n)$ by the Monotone Convergence Theorem, and

$$\begin{aligned} \sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty &\Longrightarrow \mathrm{E}[N] < \infty \\ &\Longrightarrow \mathrm{P}(N < \infty) = 1 \\ &\Longrightarrow \mathrm{P}(N = \infty) = 0 \\ &\Longrightarrow \mathrm{P}(A_n \text{ i.o.}) = 0 \qquad \text{because } (N = \infty) \equiv (A_n \text{ i.o.}). \; \square \end{aligned}$$

Pf of BCL II We will show that $\mathrm{P}(A_n^c \text{ ev.}) = 0$.

$$\begin{aligned}
\mathrm{P}(A_n^c \text{ ev.}) &= \lim_{n\to\infty} \mathrm{P}(\cap_{m\geq n} A_m^c) \;=\; \lim_{n\to\infty} \prod_{m\geq n} \mathrm{P}(A_m^c) && (1)\\
&= \lim_{n\to\infty} \prod_{m\geq n} (1-\mathrm{P}(A_m)) \;\leq\; \lim_{n\to\infty} \prod_{m\geq n} \exp\left(-\mathrm{P}(A_m^c)\right) && (2)\\
&= \lim_{n\to\infty} \exp\left(-\sum_{m\geq n} \mathrm{P}(A_m^c)\right) = 0
\end{aligned}$$

For (1) we used the following fact (due to the independence of A_n);

$$\mathrm{P}(\cap_{m\geq n} A_m^c) = \lim_{N\to\infty} \mathrm{P}(\cap_{n\leq m\leq N} A_m^c) = \lim_{N\to\infty} \prod_{n\leq m\leq N} \mathrm{P}(A_m^c) = \prod_{n\leq m} \mathrm{P}(A_m^c)$$

and $1-x \leq \exp(-x)$ was used in (2). $\square$

As a trivial example, consider $A_n = (0, 1/n)$ in $(0,1)$. Then, $\mathrm{P}(A_n) = 1/n$, $\sum \mathrm{P}(A_n) = \infty$, but $\mathrm{P}(A_n \text{ i.o.}) = \mathrm{P}(\emptyset) = 0$.

Intuitive example Consider random walk in $\mathbf{Z}^d$, $d = 0, 1, \cdots$ $S_n = X_1 + \cdots + X_n,$, $n = 0, 1, \cdots$ where X_i are independent in $\mathbf{Z}^d$. In the simplest case, each X_i has uniform distribution on 2^d possible strings. i.e., if $d = 3$, we have $2^3 = 8$ neighbors

$$\left\{ \begin{array}{c} (+1,+1,+1) \\ \vdots \\ (-1,-1,-1) \end{array} \right\}.$$

Note that each coordinate of S_n does a simple coin-tossing walk independently. We can prove that

$$\mathrm{P}(S_n = 0 \text{ i.o.}) = \begin{cases} 1 & \text{if } d = 1 \text{ or } 2 \quad (\text{recurrent}) \\ 0 & \text{if } d \geq 3 \qquad\quad\;\; (\text{transient}). \end{cases} \tag{3}$$

Sketch of Pf of (3)
Let us start with $d = 1$, then

$$\mathrm{P}(S_{2n} = 0) = \binom{2n}{n} 2^{-2n} \sim \frac{c}{\sqrt{n}} \text{ as } n \longrightarrow \infty.$$

where we used the fact, $n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$.

Note

$$\sum \left(\frac{1}{\sqrt{n}}\right)^d = \begin{cases} \infty & d = 1, 2 \\ < \infty & d = 3, 4, \cdots \end{cases} \tag{4}$$

BC II and (4) together gives (3). $\square$

Because

$$X_n \longrightarrow X \text{ a.s.} \qquad \Longleftrightarrow \qquad X_n - X \longrightarrow 0 \text{ a.s.},$$

thus it is enough to understand as convergence to 0.

Proposition 2.2 *The following are equivalent:*

1. $X_n \longrightarrow 0$
2. $\forall \epsilon > 0, \mathrm{P}(| X_n |> 0 \; i.o.) = 0$
3. $M_n \longrightarrow 0$ *where* $M_n := \sup_{n \leq k} | X_k |$
4. $\exists \epsilon_n \downarrow 0 \; : \; \mathrm{P}(| X_n |> \epsilon_n \; i.o.) = 0$

If we need to show $X_n \xrightarrow{\text{a.s.}} X$ but do not know X, then it might be easier to show instead that $\mathrm{P}(X_n \text{ is a Cauchy sequence}) = 1$. This leads to the following;

Lemma 2.3 *Let X_n be any sequence of random variables, and define $M_n := \sup_{n \leq m} | X_n - X_m |$. Then*

$$\exists X \; : \; X_n \longrightarrow X \; a.s. \quad \Longleftrightarrow \quad M_n \xrightarrow{p} 0$$

Proof Consider $M_n^* := \sup_{n \leq m,p} | X_m - X_p |$. Notice $M_n^* \downarrow$. Thus

$$M_n^* \xrightarrow{p} 0 \qquad \Longleftrightarrow \qquad M_n^* \xrightarrow{\text{a.s.}} 0$$

Combine with the previous result to finish the proof. $\square$