Stat205A: Probability Theory (Fall 2002) **Lecture: 8**

Basic $\mathcal{L}^2$ Convergence Theorem and Kolmogorov's Law of Large Numbers

Lecturer: James W. Pitman *Scribe: Vinod Prabhakaran* vinodmp@eecs.berkeley.edu

Theorem 8.1 (Basic $\mathcal{L}^2$ Convergence Theorem). [2, p. 63, (8.3)] Let X_1 X_2, ... be independent random variables with $E(X_i) = 0$ and $E(X_i^2) = \sigma_i^2 < \infty$, $i = 1, 2, \ldots$, and $S_n = X_1 + X_2 + \cdots + X_n$. If $\sum_{i=1}^{\infty} \sigma_i^2 < \infty$, then S_n converges a.s. and in $\mathcal{L}^2$ to some S_∞ with $E(S_\infty^2) = \sum_{i=1}^{\infty} \sigma_i^2$

Proof. First note that $\mathcal{L}^2$ convergence and existence of S_∞ is implied by orthogonality of X_i's: $E(X_iX_j) = 0$, $i \neq j$

$$E(S_n^2) = \sum_{i=1}^{n} \sigma_i^2$$

$$E((S_n - S_m)^2) = \sum_{i=m+1}^{n} \sigma_i^2 \rightarrow 0 \text{ as } m, n \rightarrow \infty$$

$\therefore$ S_n is Cauchy in $\mathcal{L}^2$. Since $\mathcal{L}^2$ is complete, there is a unique S_∞ (up to a.s. equivalence) such that $S_n \rightarrow S_\infty$ in $\mathcal{L}^2$.

Turning to a.s convergence, the method is to show the sequence (S_n) is a.s. Cauchy. The limit of S_n then exists a.s. by completeness of the set of real numbers. The same argument applies more generally to martingale differences X_i [2, p. 252 (4.5) for $p = 2$]. Note that this method gives S_∞ more explicitly, and does not appeal to completeness of $\mathcal{L}^2$.

Recall that S_n is Cauchy a.s. means $M_n := \sup_{p,q \geq n} |S_p - S_q| \rightarrow 0$ a.s. Note that $0 \leq M_n(\omega) \downarrow$ implies that $M_n(\omega)$ converges to a limit in $[0, \infty]$. So, if $P(M_n > \epsilon) \rightarrow 0$, $\forall\, \epsilon > 0$, then $M_n \downarrow 0$ a.s.

Let $M_n^* := \sup_{p \geq n} |S_p - S_n|$. Since by triangle inequality,

$$|S_p - S_q| \leq |S_p - S_n| + |S_q - S_n| \Rightarrow M_n^* \leq M_n \leq 2M_n^*$$

it is sufficient to show that $M_n^* \xrightarrow{P} 0$

For all $\epsilon > 0$,

$$\begin{aligned} P\left(\sup_{p \geq n} |S_p - S_n| > \epsilon\right) &= \lim_{N \to \infty} P\left(\max_{n \leq p \leq N} |S_p - S_n| > \epsilon\right) \\ &\leq \lim_{N \to \infty} \sum_{i=n+1}^{N} \frac{\sigma_i^2}{\epsilon^2} = \sum_{i=n+1}^{\infty} \frac{\sigma_i^2}{\epsilon^2} \end{aligned}$$

where the inequality is Kolmogorov's [2, p. 62, (8.2)].
Since $\sum_{i=1}^{\infty} \sigma_i^2 < \infty$,

$$\lim_{n \to \infty} P\left(\sup_{p \leq n} |S_p - S_n| > \epsilon\right) = 0$$

□

Remark: Just orthogonality rather than independence of the X_i is not enough to get an a.s. limit. Counter examples are hard [1]. According to a classical results of Rademacher-Menchoff [3, Theorems 2.3.2 and 2.3.3], for orthogonal X_i the condition $\sum_i (\log^2 i)\sigma_i^2 < \infty$ is enough for a.s. convergence of S_n, whereas if $b_i \uparrow$ with $b_i = o(\log^2 i)$ there exist orthogonal X_i such that $\sum_i b_i \sigma_i^2 < \infty$ and S_n diverges almost surely.

An easy consequence of the Basic $\mathcal{L}^2$ Convergence Theorem is the sufficiency part of Kolmogorov's three-series theorem:

Theorem 8.2 (Kolmogorov). [2, p.64, (8.4)] Let $X_1, X_2, \ldots$ be independent. Fix $b > 0$. Convergence of the following three series

- $\sum_n P(|X_n| > b) < \infty$
- $\sum_n E(X_n 1_{(|X_n|<b)})$ converges to a finite limit
- $\sum_n \mathrm{var}(X_n 1_{(|X_n|<b)}) < \infty$

is equivalent to $P(\sum_n X_n$ converges to a finite limit$) = 1$

Note. If any one of the three series diverges then

$$P\left(\sum_n X_n \text{ converges to a finite limit}\right) = 0$$

by Kolmogorov's zero-one law [2, p. 62, (8.1)]. Note also that if one or more of the series diverges for some b, then one or more of the series must diverge for every b, but exactly which of the three series diverge may depend on b. Examples can be given of 8 possible combinations of convergence/divergence.

Proof of sufficiency. That is, convergence of all 3 series implies $\sum_n X_n$ converges a.s.. Let $X_n' = X_n 1_{(|X_n|\le b)}$. Since $\sum_n P(X_n' \neq X_n) = \sum_n P(|X_n| > b) < \infty$, Borel-Cantelli lemma gives $P(X_n' \neq X_n \text{ i.o.}) = 0$ which implies $P(X_n' = X_n \text{ ev.}) = 1$. Also if $X_n'(\omega) = X_n(\omega)$ ev., then $\sum_n X_n(\omega)$ converges $\Leftrightarrow \sum_n X_n'(\omega)$ converges.

$\therefore$ it is enough to show that

$$P\left(\sum_n X_n' \text{ converges to a finite limit}\right) = 1$$

Now

$$\sum_{n=1}^N X_n' = \sum_{n=1}^N (X_n' - E(X_n')) \;+\; \sum_{n=1}^N E(X_n').$$

$\sum_{n=1}^N E(X_n')$ has a limit as $N \to \infty$ by hypothesis, and

$$\sum_{n=1}^\infty E((X_n' - E(X_n'))^2) = \sum_{n=1}^\infty \mathrm{var}(X_n') < \infty$$

implies that $\sum_{n=1}^\infty (X_n' - E(X_n'))$ converges a.s. by the basic $\mathcal{L}^2$ convergence theorem. For proof of the converse, see [2, p. 118, Example 4.7]. □

Recall *Kronecker's lemma* [2, p. 64, (8.5)]: If $a_n \uparrow \infty$ and $\sum_{n=1}^\infty X_n/a_n$ converges a.s., then $(\sum_{m=1}^n X_m)/a_n \to 0$ a.s.

Let $X_1, X_2, \ldots$ be independent with mean 0 and $S_n = X_1 + X_2 + \cdots + X_n$. If $\sum_{n=1}^\infty E(X_n^2)/a_n^2 < \infty$, then by basic $\mathcal{L}^2$ convergence theorem, $\sum_{n=1}^\infty X_n/a_n$ converges a.s. Then $S_n/a_n \to 0$ a.s.

Example. Let $X_1, X_2, \ldots$ be i.i.d., $E(X_i) = 0$, and $E(X_i^2) = \sigma^2 < \infty$.
Take $a_n = n$,

$$\sum_{n=1}^{\infty} \frac{\sigma^2}{n^2} < \infty \;\Rightarrow\; \frac{S_n}{n} \overset{a.s.}{\to} 0$$

Now take $a_n = n^{\frac{1}{2}+\epsilon}$, $\epsilon > 0$

$$\sum_{n=1}^{\infty} \frac{\sigma^2}{n^{1+2\epsilon}} < \infty \;\Rightarrow\; \frac{S_n}{n^{\frac{1}{2}+\epsilon}} \overset{a.s.}{\to} 0$$

The definitive result of this kind is the *Law of the iterated logarithm* [2, p. 434].

Theorem 8.3 (Kolmogorov's Law of Large Numbers). Let $X, X_1, X_2, \ldots$ be i.i.d. with $E(|X|) < \infty$. Let $S_n = X_1 + X_2 + \cdots + X_n$, then $S_n/n \to E(X)$ a.s. as $n \to \infty$

Note. The theorem is true with just pairwise independence instead of the full independence assumed here [2, p. 56 (7.1)]. The theorem also has an important generalization to stationary sequences (The Ergodic Theorem [2, p. 341]).

Proof. Step 1: Without loss of generality, we can assume $E(X) = 0$.

Step 2: *Truncated variables*
Define

$$\hat{X}_n := X_n 1_{(|X_n| \leq n)}$$

Note that $\hat{X}_n$ are independent. Define their centered versions $\tilde{X}_n := \hat{X}_n - E(\hat{X}_n)$
Plan: We will show that

$$\left(\frac{S_n}{n} \to 0\right) \underset{\text{(a)}}{\overset{a.s.}{=}} \left(\frac{\hat{S}_n}{n} \to 0\right) \underset{\text{(b)}}{\overset{a.s.}{=}} \left(\frac{\tilde{S}_n}{n} \to 0\right),$$

where $\hat{S}_n = \hat{X}_1 + \hat{X}_2 + \cdots + \hat{X}_n$ and $\tilde{S}_n = \tilde{X}_1 + \tilde{X}_2 + \cdots + \tilde{X}_n$.
Then using Kronecker's lemma we will show that $P\left(\tilde{S}_n/n \to 0\right) = 1$.

(a) $P(X_n = \hat{X}_n \; ev.) = 1$ because $P(X_n \neq \hat{X}_n \; i.o.) = 0$ which follows from

$$\sum_{n=1}^{\infty} P(X_n \neq \hat{X}_n) = \sum_{n=1}^{\infty} P(|X_n| > n) = \sum_{n=1}^{\infty} P(|X| > n) \leq E(|X|) < \infty$$

by Borel-Cantelli lemma. So, S_n and $\hat{S}_n$ differ only at a finite number of terms.

$$\therefore \left(\frac{S_n}{n} \to 0\right) \overset{a.s.}{=} \left(\frac{\hat{S}_n}{n} \to 0\right)$$

(b)

$$\frac{\tilde{S}_n - \hat{S}_n}{n} = \frac{1}{n} \sum_{m=1}^{n} E(\hat{X}_m) \to 0$$

since

$$E(\hat{X}_n) = E(X 1_{(|X| \leq n)}) \to E(X) = 0$$

by dominated convergenece theorem. (Dominate by $|X|$ and note $E(|X|) < \infty$.)

To finish, by Kronecker's lemma and basic $\mathcal{L}^2$ convergence theorem, it is enough to show that

$$\sum_{n=1}^{\infty} \frac{E(\tilde{X}_n^2)}{n^2} < \infty.$$

$$E(\tilde{X}_n^2) = \mathrm{var}(\hat{X}_n) \leq E(\hat{X}_n^2) = E(X^2 1_{(|X| \leq n)})$$

But, (a fact about real numbers)[1]

$$\sum_{n=1}^{\infty} \frac{X^2 1_{(|X| \leq n)}}{n^2} \leq 2|X|$$

Take expectations to complete the proof. □

References

[1] G. Alexitz. *Convergence problems of orthogonal series.* Pergamon, Oxford, 1961.

[2] R. Durrett. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition, 1996.

[3] William F. Stout. *Almost sure convergence.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Probability and Mathematical Statistics, Vol. 24.

[1] This fact can be shown roughly as follows

$$\sum_{n=1}^{\infty} \frac{x^2 1_{(|x| \leq n)}}{n^2} \approxeq x^2 \sum_{n=|x|}^{\infty} \frac{1}{n^2} \approxeq x^2 \frac{1}{|x|} \approxeq |x|$$

.