Lecture 10 : Kolmogorov's Law of Large Numbers

STAT205 *Lecturer: Jim Pitman* Scribe: *YiFeng Li* <yfli@berkeley.edu>

(These notes are a revision of the work of Vinod Prabhakaran, 2002.)

10.1 Law of the Iterated Logarithm

Let X_1, X_2, ... be i.i.d. with $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = \sigma^2$, $S_n = X_1 + ... + X_n$. We know

$$\frac{S_n}{n^{\frac{1}{2}+\varepsilon}} \xrightarrow{a.s} 0 \quad \text{as } n \to \infty.$$

We will show later

$$\frac{S_n}{\sigma n^{\frac{1}{2}}} \xrightarrow{d} N(0,1) \text{ as } n \to \infty.$$

For general interest, we state, without proof, the *Law of the Iterated Logarithm*:

$$\limsup_{n\to\infty} \frac{S_n}{\sigma\sqrt{2n\log(\log n)}} = 1 \text{ a.s.}$$

$$\liminf_{n\to\infty} \frac{S_n}{\sigma\sqrt{2n\log(\log n)}} = 1 \text{ a.s.}$$

$$\mathbb{P}(S_n > (1+\varepsilon)\sigma\sqrt{2n\log(\log n)} \text{ i.o.}) = 0$$

$$\mathbb{P}(S_n > (1-\varepsilon)\sigma\sqrt{2n\log(\log n)} \text{ i.o.}) = 0$$

10.2 Kolmogorov's Law of Large Numbers

Theorem 10.1 *Let X_1, X_2, ... be i.i.d. with $\mathbb{E}(|X_i|) < \infty$, $S_n = X_1 + ... + X_n$. Then $S_n/n \to \mathbb{E}(X)$ a.s. as $n \to \infty$.*

Note that the theorem is true with just pairwise independence instead of the full independence assumed here [[1], p.55 (7.1)]. The theorem also has an important generaliztion to stationary sequences (the *ergodic theorem*, [[1], p.337 (2.1)]).

Proof: *Step 1*: Replace X_i by $\widetilde{X}_i = X_i - \mathbb{E}X$ (note $\mathbb{E}X_i = \mathbb{E}X$). Then

$$\frac{\widetilde{S}_n}{n} = \frac{S_n}{n} - \mathbb{E}X.$$

So it's enough to consider $\mathbb{E}X = 0$.

Step 2: Now we assume $\mathbb{E}X = 0$. Introduce truncated variables

$$\widehat{X}_n := X_n I\left(|X_n| \leq n\right).$$

Observe that

$$\mathbb{P}(X_n = \widehat{X}_n \ ev.) = 1.$$

(To see this, check

$$\mathbb{P}(X_n \neq \widehat{X}_n \ i.o.) = \mathbb{P}(|X_n| > n \ i.o.)$$

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > n\ ) = \sum_{n=1}^{\infty} \mathbb{P}(|X| > n\ ) = \mathbb{E}\left(\sum_{n=1}^{\infty} I\left(|X| > n\right)\right) < \infty$$

since

$$\sum_{n=1}^{\infty} I\left(|X| > n\right) = \sum_{1 \leq n < X} 1 \leq |x| + 1.$$

Compare this to the tail sum formula for a random variable$X$ with values in $0, 1, 2, ...$

$$\mathbb{E}X = \sum_{n=1}^{\infty} n\mathbb{P}(X = n) = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n).$$

Step 3: Center the truncated variables. Define $\widetilde{X}_n := \widehat{X}_n - \mathbb{E}\left(\widehat{X}_n\right)$.
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 $\mathbb{P}\left(\tilde{S}_n/n \to 0\right) = 1$.

(a) comes from the fact that if $\omega \in \left\{\omega : X_n\left(\omega\right) = \widehat{X}_n\left(\omega\right) \ ev.\right\}$ (which has probablity 1), then $S_n\left(\omega\right) - \widehat{S}_n\left(\omega\right)$ is eventually not dependent on n. So

$$\frac{S_n\left(\omega\right) - \widehat{S}_n\left(\omega\right)}{n} \to 0 \ \text{ for such } \omega.$$

(b) comes from

$$\frac{\widehat{S}_n}{n} - \frac{\widetilde{S}_n}{n} = \frac{\mathbb{E}\widehat{X}_1 + \mathbb{E}\widehat{X}_2 + ... + \mathbb{E}\widehat{X}_n}{n} \to 0 \text{ as } n \to \infty \quad \text{(By analysis and } \mathbb{E}\widehat{X}_i \to 0\text{)}$$

But

$$\mathbb{E}\widehat{X}_n = \mathbb{E}[X_n I\left(|X_n| \le n\right)] = \mathbb{E}[X I\left(|X| \le n\right)] \to \mathbb{E}X \text{ as } n \to \infty.$$

as the integrand is dominated by $|X|$ and note
$E(|X|) < \infty$.

Now, we use Kronecker's lemma and the $\mathcal{L}^2$ convergence theorem to show that

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left(\widetilde{X}_n^2\right)}{n^2} < \infty.$$

$$\mathbb{E}\left(\widetilde{X}_n^2\right) = \mathbb{E}\left[\left(\widehat{X}_n - \mathbb{E}\left(\widehat{X}_n\right)\right)^2\right] = \mathbb{E}\left[\left(X\mathbf{1}\left(|X| \le n\right) - \mathbb{E}(X\mathbf{1}\left(|X| \le n\right))\right)^2\right]$$
$$\le \mathbb{E}\left(X\mathbf{1}\left(|X| \le n\right)\right)^2.$$

So

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left(\widetilde{X}_n^2\right)}{n^2} \le \sum_{n=1}^{\infty} \frac{\mathbb{E}X^2 I\left(|X| \le n\right)}{n^2} = \mathbb{E}\left(X^2 \sum_{n=1}^{\infty} \frac{I\left(|X| \le n\right)}{n^2}\right)$$
$$\approxeq \mathbb{E}\left(\frac{X^2}{|X|}\right) = \mathbb{E}(|X|) < \infty.$$

This came from

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

■

10.3 Convergence in distribution

Definition 10.2 $X_n \xrightarrow{d} X$ *if* $\mathbb{P}(X_n \le x) \to \mathbb{P}(X \le x)$ *for all* x *at which* $x \to \mathbb{P}(X \le x)$ *is continuous. We call this* convergence in distribution *or* weak convergence.

Note. This is really a notion of convergence of probability measures rather than of convergence of random variables. Now the limit random variable X is only unique in distribution, not unique almost surely. Obviously, if $X \stackrel{d}{=} Y$ and $X_n \xrightarrow{d} X$, then $X_n \xrightarrow{d} Y$.

Theorem 10.3 (Skorokhod) $X_n \xrightarrow{d} X \iff$ *there exists a probability with space random variables* Y_n *with* $Y_n \stackrel{d}{=} X_n$, $Y \stackrel{d}{=} X$ *and* $Y_n \xrightarrow{a.s.} Y$.

Proof: Take a single uniform variable U and use it to create the Y_n and Y. Let

$$F_n(x) = \mathbb{P}(X_n \le x)\,,$$

$$F(x) = \mathbb{P}(X \le x)\,,$$

$$Y_n = F_n^{-1}(U),$$

$$Y = F^{-1}(U).$$

Check that

$$F^{-1}(U) = \inf\{x : f(x) > u\}\,.$$

■

The following proposition is an application of the above.

Proposition 10.4 *If* $X_n \xrightarrow{d} X$*, then for every bounded continuous function* $f : R \to R$

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]\,.$$

Proof: Without loss of generality, $X_n \xrightarrow{a.s.} X$. Then $f(X_n) \to f(X)$ is bounded, so we can take expectations and use the bounded convergence theorem. ■

References

[1] Richard Durrett. *Probability: theory and examples, 3rd edition.* Thomson Brooks/Cole, 2005.