Lecture 16

Instructor: Quan Zhou

For more details about the materials covered in this note, see Chapter 4.5 of Resnick [2] and Chapters 2.3 abd 2.4 of Durrett [1].

16.1 Borel-Cantelli lemmas

Theorem 16.1 (Borel-Cantelli lemma). *If* $\sum_{n=1}^{\infty} \mathsf{P}(A_n) < \infty$*, then*[1]

$$\mathsf{P}(A_n,\ i.o.) = \mathsf{P}(\limsup_{n\to\infty} A_n) = 0.$$

Proof. Recall that $\limsup_{n\to\infty} A_n = \{\omega\colon \sum \mathbb{1}_{A_n}(\omega) = \infty\}$. Hence, by letting $N(\omega) = \sum_{n=1}^{\infty} \mathbb{1}_{A_n}(\omega)$, we only need to show $N(\omega) < \infty$ with probability 1. By MCT,

$$E[N] = \int \sum_{n=1}^{\infty} \mathbb{1}_{A_n} d\mathsf{P} = \sum_{n=1}^{\infty} \int \mathbb{1}_{A_n} d\mathsf{P} = \sum_{n=1}^{\infty} \mathsf{P}(A_n) < \infty,$$

which implies that $\mathsf{P}(N < \infty) = 1$. □

Theorem 16.2 (Second Borel-Cantelli lemma). *If the events A_n are independent, then* $\sum_{n=1}^{\infty} \mathsf{P}(A_n) = \infty$ *implies* $\mathsf{P}(A_n,\ i.o.) = \mathsf{P}(\limsup_{n\to\infty} A_n) = 1$.

Proof. Let $M < N < \infty$. Using $e^x > 1 + x$ we get

$$\mathsf{P}\left(\cap_{n=M}^{N} A_n^c\right) = \prod_{n=M}^{N} (1 - \mathsf{P}(A_n)) \leq \prod_{n=M}^{N} \exp(-\mathsf{P}(A_n)) \to 0, \quad \text{as } N \to \infty.$$

Since $\left(\cap_{n=M}^{N} A_n^c\right)^c = \cup_{n=M}^{N} A_n$, we get $\mathsf{P}(\cup_{n=M}^{\infty} A_n) = 1$, which is true for every M. Thus, by the continuity of probability measures,

$$\mathsf{P}(\limsup_{n\to\infty} A_n) = \mathsf{P}\left(\lim_{M\to\infty} (\cup_{n=M}^{\infty} A_n)\right) = \lim_{M\to\infty} \mathsf{P}(\cup_{n=M}^{\infty} A_n) = 1,$$

which completes the proof. □

[1] "i.o." means infinitely often.

Example 16.1. Consider a sequence of random variables $X_1, X_2, \dots$ such that $\mathsf{P}(X_n = n) = 1/n$ and $\mathsf{P}(X_n = 0) = 1 - 1/n$. Recall that $X_n \xrightarrow{P} 0$. If we further assume $X_1, X_2, \dots$ are independent, then by the second Borel-Cantelli lemma, they do not converge almost surely, since

$$\sum_{n=1}^{\infty} \mathsf{P}(X_n \geq 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Example 16.2. Let $X_1, X_2, \dots$ be i.i.d. exponential random variables with density function $f(x) = e^{-x}$ on $[0, \infty)$. Note that $\mathsf{P}(X_n / \log n > \epsilon) = n^{-\epsilon} \to 0$ for every $\epsilon > 0$. Hence, $X_n / \log n \xrightarrow{P} 0$. However, when $\epsilon \in (0, 1)$, by the two Borel-Cantelli lemmas, we have $\mathsf{P}(X_n / \log n > 1 - \epsilon, \text{ i.o.}) = 1$ and $\mathsf{P}(X_n / \log n > 1 + \epsilon, \text{ i.o.}) = 0$. This further implies, by choosing a sequence $\epsilon_k \downarrow 0$, that $\mathsf{P}(\limsup_{n\to\infty} X_n / \log n = 1) = 1$.

Proposition 16.1. *Let $X_1, X_2, \dots$ be i.i.d. with $E|X_i| = \infty$, and $S_n = X_1 + \cdots + X_n$. Then $\mathsf{P}(\lim_{n\to\infty} S_n/n$ exists and is finite$) = 0$.*

Proof. Observe that for any $x \in [n, n+1]$, we have $\mathsf{P}(|X_1| > x) \leq \mathsf{P}(X_1 > n)$. Hence,

$$\infty = E|X_1| = \int_0^{\infty} \mathsf{P}(|X_1| > x) dx \leq \sum_{n=0}^{\infty} \mathsf{P}(|X_1| > n).$$

(This is a useful trick. An integral of a monotone function can be bounded from above and from below by the summation.) Therefore, by the second Borel-Cantelli lemma, $\mathsf{P}(|X_n| > n, \text{ i.o.}) = 1$. Next, we study the increment $|S_{n+1}/(n+1) - S_n/n|$. Note that

$$D_n = \frac{S_{n+1}}{n+1} - \frac{S_n}{n} - = \frac{n(S_{n+1} - S_n) - S_n}{n(n+1)} = \frac{X_{n+1}}{n+1} - \frac{S_n}{n(n+1)}.$$

Let $C = \{\omega \colon S_n/n$ converges to a finite limit$\}$ and $A = \{\omega \colon |X_n| > n, \text{ i.o.}\}$. We prove by contradiction that $A \cap C = \emptyset$. Suppose that $\omega \in A \cap C$. we clearly have $S_n/n(n+1) \to 0$ and thus there exists $N < \infty$ such that for all $n \geq N$, we have $|S_n/n(n+1)| < \epsilon \in (0, 1)$. But $\omega \in A$ also implies that

$$\sum_{n=N+1}^{\infty} \mathbb{1}_{\{|X_n| > n\}}(\omega) = \sum_{n=N+1}^{\infty} \mathbb{1}_{\{|X_n|/n > 1\}}(\omega) = \infty.$$

For $n > N$, $|X_n|/n > 1$ yields $|D_{n-1}| \geq 1-\epsilon$. Therefore, $|D_n| \geq 1-\epsilon$ infinitely often, which by Cauchy criterion implies that S_n/n does not converge. So, this yields the contradiction and we conclude that $\mathsf{P}(A \cap C) = 0$. By the union bound, $0 = \mathsf{P}(A \cap C) \geq \mathsf{P}(A) + \mathsf{P}(C) - 1 = \mathsf{P}(C)$. □

Theorem 16.3 (SLLN with finite fourth moments)**.** *Let $X_1, X_2, \dots$ be i.i.d. with mean μ and $E[X_1^4] < \infty$, and $S_n = \sum_{i=1}^n X_i$. Then $S_n/n \overset{a.s.}{\to} 0$.*

Proof. Without loss of generality, assume $\mu = 0$.

$$E[S_n^4] = E \sum_{1 \leq i,j,k,l \leq n} X_i X_j X_k X_l.$$

Due to independence and the assumption $E[X_i] = 0$, we have

$$E[S_n^4] = nE(X_1^4) + 3n(n-1)\left(E[X_1^2]\right)^2 \leq Cn^2$$

for some $C \in (0, \infty)$. (Check it!) Hence, by Markov's inequality,

$$\mathsf{P}(|S_n|/n \geq \epsilon) \leq \frac{1}{n^4 \epsilon^4} E S_n^4 < \frac{C}{n^2 \epsilon^4}.$$

The series $\sum_{n=1}^\infty n^{-2}$ converges and thus by the Borel-Cantelli lemma,

$$\mathsf{P}(|S_n|/n \geq \epsilon \text{ i.o.}) = 0.$$

Now let's choose $\epsilon_k = 1/k$ for $k = 1, 2, \dots$. Then, by the countable sub-additivity of measures,

$$\mathsf{P}(\cup_{k \geq 1} \{|S_n|/n \geq \epsilon_k \text{ i.o.}\}) = 0,$$

which implies $\mathsf{P}(A) = 1$ where

$$A = \bigcap_{k \geq 1} \{|S_n|/n \geq \epsilon_k \text{ finitely often}\}.$$

This gives the almost sure convergence we want to prove. To see this, fix an arbitrary $\omega \in A$. The definition of A means that for any $k \geq 1$, we can find $N(\epsilon_k, \omega) < \infty$ such that for any $n \geq N(\epsilon_k, \omega)$, we have $|S_n(\omega)|/n < \epsilon_k$; that is, $|S_n(\omega)|/n \to 0$. □

16.2 Strong law of large numbers

Theorem 16.4 (Etemadi's SLLN)**.** *Let $X_1, X_2, \dots$ be pairwise independent identically distributed random variables with $E|X_1| < \infty$ and mean μ. Let $S_n = \sum_{i=1}^n X_i$. Then $S_n/n \overset{a.s.}{\to} \mu$.*

Proof. See the next section. □

Remark 16.1. If we replace "pairwise independence" with "mutual independence", this result is known as Kolmogorov's SLLN. Of course, Etemadi's result is better. Etemadi's paper was published in 1981.

Corollary 16.1. *Let $X_1, X_2, \dots$ be i.i.d. with $E(X_i^+) = \infty$ and $E(X_i^-) < \infty$. Let $S_n = X_1 + \cdots + X_n$. We have $S_n/n \to \infty$ almost surely.*

Proof. Choose an arbitrary $m > 0$, and let $Y_i(m) = X_i \wedge m$. Clearly, $Y_1(m), Y_2(m), \dots$ are still i.i.d. and $E|Y_1(m)| < \infty$. Let $T_n(m) = \sum_{i=1}^n Y_i(m)$. Then by SLLN, T_n/n converges to EY_1 a.s. Since $X_i \geq Y_i$ for each i,

$$\liminf_{n\to\infty} \frac{S_n}{n} \geq \liminf_{n\to\infty} \frac{T_n(m)}{n} = \lim_{n\to\infty} \frac{T_n(m)}{n} = E[Y_1(m)].$$

By MCT, $\lim_{m\to\infty} E[Y_1(m)] = E[\lim_{m\to\infty} Y_1(m)] = E[X_1] = \infty$. That is, $\liminf_{n\to\infty} S_n/n \geq \infty$, which yields the asserted result. □

Example 16.3 (Glivenko-Cantelli theorem)**.** Let $X_1, X_2, \dots$ be i.i.d. with distribution function F and let

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_m \leq x\}},$$

which is called the empirical distribution function. Let x be fixed and $Y_n = \mathbb{1}_{\{X_n \leq x\}}$. Clearly, $Y_1, Y_2, \dots, Y_n$ are also i.i.d. with expectation $F(x)$. SLLN implies that $(\sum_{i=1}^n Y_i)/n \overset{a.s.}{\to} F(x)$. Actually, with some extra work, one can show that as $n \to \infty$,

$$\sup_x |F_n(x) - F(x)| \overset{a.s.}{\to} 0.$$

16.3 Proof of Etemadi's SLLN

We first prove three lemmas.

Lemma 16.1. *Let* $Y_k = X_k \mathbb{1}_{\{|X_k| \leq k\}}$ *and* $T_n = Y_1 + \cdots + Y_n$. *Etemadi's SLLN would hold if* $T_n/n \overset{a.s.}{\to} \mu$.

Proof. Since $\sum_{k=1}^{\infty} \mathsf{P}(|X_k| > k) \leq \int_0^\infty \mathsf{P}(|X_1| > t)dt = E|X_1| < \infty$, we have $\mathsf{P}(X_k \neq Y_k, \text{ i.o.}) = 0$ by Borel-Cantelli lemma. Hence, almost surely there exists a finite upper bound for $|S_n - T_n|$. □

Lemma 16.2. *If* $y \geq 0$, *then* $2y \sum_{k>y} k^{-2} \leq 4$.

Proof. If $m \geq 2$, then

$$\sum_{k \geq m} \frac{1}{k^2} \leq \int_{m-1}^{\infty} \frac{1}{x^2} dx = \frac{1}{m-1}.$$

If $y \geq 1$, then the smallest integer $k > y$ is 2 and thus

$$2y \sum_{k>y} \frac{1}{k^2} \leq \frac{2y}{\lfloor y \rfloor} < 4.$$

For $0 \leq y < 1$, we have

$$2y \sum_{k>y} \frac{1}{k^2} \leq 2 + 2y \sum_{k=2}^{\infty} \frac{1}{k^2} < 2 \left(1 + \sum_{k=2}^{\infty} \frac{1}{k^2} \right) \leq 4.$$

Recall that $\sum_{n=1}^{\infty} n^{-2} = \pi^2/6$. □

Lemma 16.3. $\sum_{k=1}^{\infty} \mathrm{Var}(Y_k)/k^2 \leq 4E|X_1| < \infty$.

Proof. Note that $\mathrm{Var}(Y_k) \leq \int_0^k 2y\mathsf{P}(|X_1| > y)dy$ (recall how we proved Feller's WLLN). Hence, by Fubini's theorem and Lemma 16.2,

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{\mathrm{Var}(Y_k)}{k^2} &\leq \sum_{k=1}^{\infty} \frac{1}{k^2} \int_0^\infty \mathbb{1}_{(-\infty,k)}(y)\, 2y\, \mathsf{P}(|X_1| > y)dy \\
&= \int_0^\infty \left\{ 2y \sum_{k=1}^{\infty} k^{-2} \mathbb{1}_{\{y<k\}} \right\} \mathsf{P}(|X_1| > y)dy \\
&\leq 4 \int_0^\infty \mathsf{P}(|X_1| > y)dy = 4E|X_1|.
\end{aligned}$$

That is, $\sum_{k=1}^{\infty} \mathrm{Var}(Y_k)/k^2 < \infty$. □

Proof of Theorem 16.4. Since $\{X_n^+\}$ and $\{X_n^-\}$ both satisfy the assumptions of the theorem and $X_n = X_n^+ - X_n^-$, we can without loss of generality assume that $X_n \geq 0$ for all n and then prove the claim made in Lemma 16.1. Let $k(n) = \lfloor \alpha^n \rfloor$ for some $\alpha > 1$. By Markov's inequality, for $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathsf{P}\left(|T_{k(n)} - E[T_{k(n)}]| > \epsilon k(n)\right) \leq \frac{1}{\epsilon^2} \sum_{n=1}^{\infty} \frac{\mathrm{Var}(T_{k(n)})}{k(n)^2}.$$

By the pairwise independence of $X_1, X_2, \ldots$,

$$\mathrm{Var}(T_{k(n)}) = \mathrm{Var}\left(Y_1 + \cdots + Y_{k(n)}\right) = \sum_{m=1}^{k(n)} \mathrm{Var}(Y_m).$$

Then, by Fubini's theorem,

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}(T_{k(n)})}{k(n)^2} = \sum_{m=1}^{\infty} \mathrm{Var}(Y_m) \sum_{n: k(n) \geq m} \frac{1}{k(n)^2}.$$

Using $k(n) = \lfloor \alpha^n \rfloor$ and $\lfloor \alpha^n \rfloor \geq \alpha^n / 2$ for $n \geq 1$,

$$\sum_{n: k(n) \geq m} \frac{1}{k(n)^2} = \sum_{n: \lfloor \alpha^n \rfloor \geq m} \frac{1}{\lfloor \alpha^n \rfloor^2} \leq 4 \sum_{n: \alpha^n \geq m} \frac{1}{\alpha^{2n}} \leq \frac{4}{m^2 (1 - \alpha^{-2})}.$$

Hence, by Lemma 16.3,

$$\sum_{n=1}^{\infty} \mathsf{P}\left(|T_{k(n)} - E[T_{k(n)}]| > \epsilon k(n)\right) \leq \frac{4}{\epsilon^2 (1 - \alpha^{-2})} \sum_{m=1}^{\infty} \frac{\mathrm{Var}(Y_m)}{m^2} < \infty.$$

Since ϵ is arbitrary, by Borel-Cantelli lemma, we get $[T_{k(n)} - E(T_{k(n)})]/k(n) \stackrel{a.s.}{\to} 0$. By DCT, $E[Y_k] \to E[X_1]$ as $k \to \infty$ and thus $E[T_{k(n)}]/k(n) \to E[X_1]$, which yields $T_{k(n)} \stackrel{a.s.}{\to} \mu$. To establish the a.s. convergence for the entire sequence, observe that for $k(n) \leq m < k(n+1)$, we have

$$\frac{T_{k(n)}}{k(n+1)} \leq \frac{T_m}{m} \leq \frac{T_{k(n+1)}}{k(n)}$$

by the non-negativity of $\{Y_n\}$. Since $k(n+1)/k(n) \to \alpha$, using the a.s. convergence of $\{T_{k(n)}\}$ we get

$$\frac{1}{\alpha} E(X_1) \leq \liminf_{n \to \infty} \frac{T_m}{m} \leq \limsup_{n \to \infty} \frac{T_m}{m} \leq \alpha E(X_1).$$

Letting $\alpha \downarrow 1$, we get the asserted SLLN. □

References

[1] Rick Durrett. *Probability: Theory and Examples*, volume 49. Cambridge university press, 2019.

[2] Sidney Resnick. *A Probability Path*. Springer, 2019.