Lecture 17

Instructor: Quan Zhou

For more details about the materials covered in this note, see Chapters 4.5 and 7.3 to 7.6 of Resnick [2] and Chapter 2.5 of Durrett [1].

17.1 Kolmogorov zero-one law

Definition 17.1. Consider a sequence of random variables $X_1, X_2, \dots$. Let

$$\mathcal{F}'_n = \sigma(X_n, X_{n+1}, \dots) = \sigma\left(\cup_{i=n}^{\infty} \sigma(X_i)\right)$$

be the smallest σ-algebra with respect to which $X_n, X_{n+1}, \dots$ are measurable. The tail σ-algebra is defined as $\mathcal{T} = \cap_{n \geq 1} \mathcal{F}'_n$.

Example 17.1. Let B_n be a sequence of Borel sets. Then $\{X_n \in B_n, \text{ i.o.}\} \in \mathcal{T}$. In particular, by letting $X_n = \mathbb{1}_{A_n}$ and $B_n = \{1\}$, we get $\{A_n, \text{ i.o.}\} \in \mathcal{T}$.

Example 17.2. Let $S_n = X_1 + \cdots + X_n$. Then $\{S_n \text{ converges to a finite limit}\} \in \mathcal{T}$ but $\{\limsup_{n \to \infty} S_n > 0\} \notin \mathcal{T}$. However, if c_n is a sequence of real numbers such that $c_n \to \infty$, then $\{\limsup_{n \to \infty} S_n / c_n > x\} \in \mathcal{T}$.

Theorem 17.1 (Kolmogorov zero-one law)**.** *If $X_1, X_2, \dots$ are independent and $A \in \mathcal{T}$, then $\mathsf{P}(A) = 0$ or $\mathsf{P}(A) = 1$.*

Proof. We show that A is independent of itself if $A \in \mathcal{T}$.

Step 1. Recall that $\mathcal{F}_{n-1} = \sigma(X_1, \dots, X_{n-1})$ is generated by the π-system of all the events $B = \{X_1 \leq x_1, \dots, X_{n-1} \leq x_{n-1}\}$ for $x_1, \dots, x_{n-1} \in \mathbb{R}$. Further, one can use Dynkin's theorem to show that $\mathcal{F}'_n$ is generated by the π-system of all the events of the form[1]

$$C = \{X_n \leq x_n, \dots, X_{n+k} \leq x_{n+k}\}, \qquad k \in \mathbb{N},\ x_n, \dots, x_{n+k} \in \mathbb{R}.$$

Clearly, any such sets B and C are independent and thus by Lemma 7.1, $\mathcal{F}_{n-1}$ and $\mathcal{F}'_n$ are independent.

Step 2. We show that $\mathcal{F}'_1 = \sigma(X_1, X_2, \dots)$ is independent of $\mathcal{T}$. First, assume $\tilde{A} \in \mathcal{F}_k$ for some k. Since $\mathcal{T} \subset \mathcal{F}'_{k+1}$, then from Step 1 we know that

[1] In general, given σ-algebras $\mathcal{A}_1, \mathcal{A}_2, \dots$, then $\sigma(\cup_{i=1}^{\infty} \mathcal{A}_i) = \sigma(\mathcal{P})$ where $\mathcal{P}$ is the π-system $\{\cap_{i \in I} A_i : I \subset \{1, 2, \dots\}, A_i \in \mathcal{A}_i,\ 1 \leq |I| < \infty\}$.

any event in $\mathcal{T}$ is independent of $\tilde{A}$, which implies $\mathcal{T}$ and $\mathcal{F}_k$ are independent since $\tilde{A}$ is arbitrary. Since both $\cup_{n\geq 1}\mathcal{F}_n$ and $\mathcal{T}$ are π-systems, by Lemma 7.1 we have $\mathcal{T}$ and $\sigma(X_1, X_2, \dots)$ are independent.

Hence, we conclude that A is independent of itself, i.e. $\mathsf{P}(A \cap A) = \mathsf{P}(A)\mathsf{P}(A)$, which yields $\mathsf{P}(A) = 0$ or 1. □

Corollary 17.1. *Let $\mathcal{T}$ be a tail σ-algebra generated by a sequence of independent random variables. If a random variable Y is measurable w.r.t. $\mathcal{T}$, i.e. $Y \in \mathcal{T}$, then $\mathsf{P}(Y = c) = 1$ for some $c \in \mathbb{R}$.*

Proof. Let F_Y be the distribution function of Y. Observe that $F_Y(y) \in \{0, 1\}$ for all $y \in \mathbb{R}$ by Kolmogorov zero-one law. □

Example 17.3. Consider a sequence of independent random variables $\{X_n\}_{n\geq 1}$. Both $\liminf_{n\to\infty} X_n$ and $\limsup_{n\to\infty} X_n$ are measurable w.r.t. $\mathcal{T}$. Therefore, with probability 1, $\liminf_{n\to\infty} X_n$ and $\limsup_{n\to\infty} X_n$ are constants (note they may be $\pm\infty$).

17.2 Convergence of random series

We say a (random) series converges if and only if it converges to a finite limit.

Theorem 17.2 (Lévy's theorem)**.** *Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables, then $\sum_{n=1}^{\infty} X_n$ converges in probability if and only if $\sum_{n=1}^{\infty} X_n$ converges almost surely.*

Proof. See Resnick [2, Theorem 7.3.2]. □

Theorem 17.3 (Kolmogorov's maximal inequality)**.** *Let $X_1, \dots, X_n$ be independent with $E[X_i] = 0$ and $\mathrm{Var}(X_i) < \infty$ for all i. Then,*

$$\mathsf{P}(\max_{1\leq k\leq n} |S_k| \geq x) \leq \frac{\mathrm{Var}(S_n)}{x^2}.$$

Proof. See Durrett [1, Theorem 2.5.5]. □

Remark 17.1. This result is used in Kolmogorov's proof for SLLN. Note that we need the mutual independence among $X_1, X_2, \dots$. For comparison, Chebyshev's inequality yields $\mathsf{P}(|S_n| \geq x) \leq \mathrm{Var}(S_n)/x^2$.

Theorem 17.4 (Kolmogorov convergence criterion)**.** *Suppose $X_1, X_2, \dots$ are independent with $E[X_i] = 0$ for all i and $\sum_{n=1}^\infty \mathrm{Var}(X_n) < \infty$. Then, $\sum_{n=1}^\infty X_n$ converges a.s.*

Proof. See Resnick [2, Theorem 7.3.3]. □

Theorem 17.5 (Kolmogorov's three series theorem)**.** *Let $X_1, X_2, \dots$ be independent. In order that $\sum_{n=1}^\infty X_n$ converges almost surely, it is necessary and sufficient that for some $c > 0$,*

(i) $\sum_{n=1}^\infty \mathsf{P}(|X_n| > c) < \infty$,

(ii) $\sum_{n=1}^\infty E[X_n \mathbb{1}_{\{|X_n| < c\}}]$ *converges (to a finite limit),*

(iii) $\sum_{n=1}^\infty \mathrm{Var}(X_n \mathbb{1}_{\{|X_n| < c\}}) < \infty$.

Proof of the "sufficiency" part in Theorem 17.5. Define $Y_n = X_n \mathbb{1}_{\{|X_n| < c\}}$. By Kolmogorov convergence criterion and condition (iii), $\sum_{n=1}^\infty (Y_n - E[Y_n])$ converges a.s. By condition (ii), it means that $\sum_{n=1}^\infty Y_n$ converges a.s. By Borel-Cantelli lemma and condition (i), $\mathsf{P}(X_n \neq Y_n, \text{ i.o.}) = 0$. Hence, $\sum_{n=1}^\infty X_n$ converges a.s.

For the "necessity" part, see the textbook. □

Lemma 17.1 (Kronecker's lemma)**.** *If $a_n \uparrow \infty$ and $\sum\limits_{n=1}^{\infty} (x_n/a_n)$ converges,*

$$\frac{1}{a_n} \sum_{k=1}^n x_k \to 0.$$

Proof. See Resnick [2, Lemma 7.4.1]. □

Theorem 17.6 (Kolmogorov's SLLN)**.** *Let $X_1, X_2, \dots$ be i.i.d. random variables with $E|X_1| < \infty$ and mean μ. Then $S_n/n \overset{a.s.}{\to} \mu$.*

Proof. Again let's define $Y_k = X_k \mathbb{1}_{\{|X_k| \leq k\}}$ and $T_n = Y_1 + \cdots + Y_n$ and recall that it suffices to show $T_n/n \overset{a.s.}{\to} \mu$. Let $Z_k = Y_k - E(Y_k)$. Recall that by Lemma 16.3, $\sum_{k=1}^\infty \mathrm{Var}(Z_k)/k^2 < \infty$, which together with Theorem 17.4 implies that $\sum_{k=1}^\infty Z_k/k$ converges a.s. By Kronecker's lemma, this further yields that $n^{-1}\sum_{k=1}^n Z_k \to 0$, i.e.

$$\frac{T_n}{n} - \frac{1}{n}\sum_{k=1}^n E(Y_k) \to 0, \text{ a.s.}$$

By DCT, $E(Y_k) \to \mu$ and thus $n^{-1}\sum\limits_{k=1}^{n} E(Y_k) \to \mu$. Therefore, $T_n/n \overset{a.s.}{\to} \mu$. □

17.3 Rates of convergence for LLN

Theorem 17.7. *Let $X_1, X_2, \dots$ be i.i.d. random variables with $E[X_1] = 0$ and $\mathrm{Var}(X_1) < \infty$. Then for any $\epsilon > 0$,*

$$\frac{S_n}{\sqrt{n}(\log n)^{1/2+\epsilon}} \xrightarrow{a.s.} 0.$$

Proof. Let $a_n = n^{1/2}(\log n)^{1/2+\epsilon}$ for $n \geq 2$. Observe that

$$\sum_{n=2}^{\infty} \mathrm{Var}(X_n/a_n) = \sum_{n=2}^{\infty} \frac{\sigma^2}{n(\log n)^{1+2\epsilon}} < \infty.$$

(Check it! Hint: $d \log x = x^{-1} dx$.) By Theorem 17.4, $\sum_{n=1}^{\infty} X_n/a_n$ converges a.s. Apply Kronecker's lemma to conclude the proof. □

Theorem 17.8. *Let $X_1, X_2, \dots$ be i.i.d. random variables with $E[X_1] = 0$ and $E|X_1|^p < \infty$, where $1 < p < 2$. Then $S_n/n^{1/p} \xrightarrow{a.s.} 0$.*

Proof. See Durrett [1, Theorem 2.5.12]. □

Theorem 17.9 (Law of iterated logarithm)**.** *Let $X_1, X_2, \dots$ be i.i.d. random variables with $E[X_1] = 0$ and $\mathrm{Var}(X_1) = 1$. Then,*

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

Proof. See Durrett [1, Theorem 8.5.2]. □

Remark 17.2. The law of iterated logarithm is usually treated in chapters/books on random walks or Brownian motion. We can prove that

$$\frac{S_n}{\sqrt{2n \log(\log n)}} \xrightarrow{P} 0.$$

So, $\sqrt{n \log(\log n)}$ is the rate at which the limit in probability and the almost sure limit supremum/infimum are different. Another critical rate is $\sqrt{n}$. We will see that the central limit theorem implies that $S_n/\sqrt{n}$ converges in distribution to a standard normal random variable.

References

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

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