

STATS310B – Lecture 9

Sourav Chatterjee
Scribed by Michael Howes

02/01/22

Contents

| | |
|--|----------|
| 1 Uniform integrability | 1 |
| 2 Lévy downwards convergence theorem | 2 |
| 3 De Finetti's theorem | 3 |
| 3.1 The exchangeable σ -algebra | 4 |

1 Uniform integrability

We ended last lecture with the statement of the following proposition,

Proposition 1. *Let $\{X_n\}_{n \geq 1}$ be a sequence of random variables. The sequence $\{X_n\}_{n \geq 0}$ is uniformly integrable if and only if the following both hold,*

1. $\sup_n \mathbb{E}[|X_n|] < \infty$.
2. For all $\varepsilon > 0$ there exists $\delta > 0$ such that for all A and n , if $\mathbb{P}(A) < \delta$, then $\mathbb{E}[|X_n| \mathbf{1}_A] < \varepsilon$.

Which we will now prove.

Proof. First suppose that $\{X_n\}_{n \geq 1}$ is uniformly integrable. Fix any $\varepsilon > 0$, then there exists a k such that for all n ,

$$\mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| > k\}}] < \varepsilon.$$

It follows that,

$$\mathbb{E}[|X_n|] = \mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| \leq k\}}] + \mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| > k\}}] < k + \varepsilon.$$

Thus, $\sup_n \mathbb{E}[|X_n|] < \infty$. Now let $\varepsilon > 0$ be arbitrary and fix k so that for all n ,

$$\mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| > k\}}] < \varepsilon/2.$$

Set $\delta = \frac{\varepsilon}{2k}$ and suppose $\mathbb{P}(A) < \delta$. Then

$$\begin{aligned} \mathbb{E}[|X_n| \mathbf{1}_A] &= \mathbb{E}[|X_n| \mathbf{1}_{A \cap \{|X_n| > k\}}] + \mathbb{E}[|X_n| \mathbf{1}_{A \cap \{|X_n| \leq k\}}] \\ &\leq \mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| > k\}}] + \mathbb{E}[k \mathbf{1}_A] \\ &< \varepsilon/2 + k\mathbb{P}(A) \\ &< \varepsilon/2 + k\delta \\ &= \varepsilon. \end{aligned}$$

Now conversely suppose that the two conditions hold and note that,

$$\mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| > k\}}] \geq \mathbb{E}[k \mathbf{1}_{\{|X_n| > k\}}] = k\mathbb{P}(|X_n| > k).$$

Thus, for every n and k , $\mathbb{P}(|X_n| > k) \leq \frac{\sup_m \mathbb{E}[|X_m|]}{k}$. Thus, given $\epsilon > 0$ choose δ so that $\mathbb{P}(A) < \delta$ implies that $\mathbb{E}[|X_n| \mathbf{1}_A] < \epsilon$ for every n . If $k > \frac{\sup_m \mathbb{E}[|X_m|]}{\delta}$, then $\mathbb{P}(|X_n| > k) < \epsilon$ for all n and hence

$$\mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| > k\}}] < \epsilon. \quad \square$$

The above proposition has the following important corollary.

Corollary 1. *Let X be an integrable random variable. Then for all $\epsilon > 0$, there exists $\delta > 0$ such that if $\mathbb{P}(A) < \delta$, then $\mathbb{E}[X \mathbf{1}_A] < \epsilon$.*

Proof. By proposition (1) it suffices to show that the sequence $X_n = X$ is uniformly integrable. By the dominated convergence theorem

$$\lim_{k \rightarrow \infty} \mathbb{E}[|X| \mathbf{1}_{\{|X| > k\}}] = 0,$$

showing that for all ϵ there exists a k such that $\mathbb{E}[|X| \mathbf{1}_{\{|X| > k\}}] < \epsilon$. \square

2 Lévy downwards convergence theorem

With these results about uniform integrability we are ready to finish proving Lévy downwards convergence theorem. Recall the theorem's statement,

Theorem 1 (Lévy's downwards convergence theorem). *Let X be an integrable random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{F}_0 \supseteq \mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \dots$ be a decreasing sequence of sub- σ -algebras of \mathcal{F} . Define $\mathcal{F}^* = \bigcap_{n=0}^{\infty} \mathcal{F}_n$, then*

$$\mathbb{E}(X|\mathcal{F}_n) \rightarrow \mathbb{E}(X|\mathcal{F}^*),$$

almost surely and in L^1 .

Proof. In the previous lecture we use the upcrossing lemma to show that there exists an integrable random variable X^* such that if $X_n = \mathbb{E}(X|\mathcal{F}_n)$, then $X_n \rightarrow X^*$ almost surely. It remains to show that $X_n \rightarrow X^*$ in L^1 and that $X^* = \mathbb{E}(X|\mathcal{F}^*)$. We will first show that $\{X_n\}_{n \geq 0}$ is uniformly integrable. Thus, fix $\epsilon > 0$ and let δ be such that $\mathbb{P}(A) < \delta$ implies that $\mathbb{E}[|X| \mathbf{1}_A] < \epsilon$. Now choose k so that $\frac{\mathbb{E}[|X|]}{k} < \epsilon$. Let A_n be the event $\{\mathbb{E}(|X||\mathcal{F}_n) > k\}$. Note that, by Markov's inequality,

$$\begin{aligned} \mathbb{P}(A_n) &\leq \frac{\mathbb{E}[\mathbb{E}(|X||\mathcal{F}_n)]}{k} \\ &= \frac{\mathbb{E}[|X|]}{k} \\ &< \delta. \end{aligned}$$

Thus, $\mathbb{E}[|X| \mathbf{1}_{A_n}] < \epsilon$. Furthermore, since $A_n \in \mathcal{F}_n$ we have

$$\mathbb{E}[\mathbb{E}(|X||\mathcal{F}_n) \mathbf{1}_{A_n}] = \mathbb{E}[|X| \mathbf{1}_{A_n}] < \epsilon.$$

By Jensen's inequality $|X_n| = |\mathbb{E}(X|\mathcal{F}_n)| \leq \mathbb{E}(|X||\mathcal{F}_n)$. Thus, $\mathbb{E}[|X_n| \mathbf{1}_{A_n}] < \epsilon$. Furthermore, since $|X_n| < \mathbb{E}(|X||\mathcal{F}_n)$, we have $\{|X_n| > k\} \subseteq A_n$ and so

$$\mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| > k\}}] \leq \mathbb{E}[|X_n| \mathbf{1}_{A_n}] < \epsilon.$$

Thus, $\{X_n\}_{n \geq 0}$ are uniformly integrable. Since $X_n \rightarrow X^*$ almost surely, this implies that $X_n \rightarrow X^*$ in L^1 . It remains to show that $X^* = \mathbb{E}(X|\mathcal{F}^*)$. First note that X^* is \mathcal{F}^* measurable. This is because for every m , $X_n = \mathbb{E}(X|\mathcal{F}_n)$ is \mathcal{F}_m measurable for every $n \geq m$. Thus, $X^* = \lim_{n \geq m} X_n$ is \mathcal{F}_m measurable. Since this holds for every m we must have that X^* is $\mathcal{F}^* = \bigcap_{m=1}^{\infty} \mathcal{F}_m$ measurable. Now suppose that $A \in \mathcal{F}^*$. Then $A \in \mathcal{F}_n$ for every n and hence

$$\mathbb{E}[X_n \mathbf{1}_A] = \mathbb{E}[X \mathbf{1}_A],$$

for every m . Furthermore,

$$|\mathbb{E}[X_n \mathbf{1}_A] - \mathbb{E}[X^* \mathbf{1}_A]| \leq \mathbb{E}[|X_n - X^*| \mathbf{1}_A] \leq \mathbb{E}[|X_n - X^*|].$$

Since $X_n \rightarrow X^*$ in L^1 , this implies that

$$\mathbb{E}[X^* \mathbf{1}_A] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n \mathbf{1}_A] = \lim_{n \rightarrow \infty} \mathbb{E}[X \mathbf{1}_A] = \mathbb{E}[X \mathbf{1}_A].$$

Thus, $X^* = \mathbb{E}(X|\mathcal{F}^*)$. □

We will now discuss and later prove an application of this convergence theorem.

3 De Finetti's theorem

Definition 1. Let X_1, X_2, \dots be an infinite sequence of random variables. The sequence $\{X_n\}_{n \geq 1}$ is *exchangeable* if for any permutation $\pi : \mathbb{N} \rightarrow \mathbb{N}$ that fixes all but finitely many values $(X_{\pi(1)}, X_{\pi(2)}, \dots)$ has the same distribution as (X_1, X_2, \dots) .

More precisely, we can think of the sequence $\{X_n\}_{n \geq 1}$ as a single random variable X taking values in $\mathbb{R}^{\mathbb{N}}$ given by

$$X(\omega) = (X_1(\omega), X_2(\omega), \dots).$$

The distribution of X is the probability measure it induces on $(\mathbb{R}^{\mathbb{N}}, \mathcal{F}_{\mathbb{R}^{\mathbb{N}}})$ where $\mathcal{F}_{\mathbb{R}^{\mathbb{N}}}$ is the product σ -algebra of countably many copies of the Borel σ -algebra. Thus, the law of X is the probability measure μ_X given by

$$\mu_X(B) = \mathbb{P}(X^{-1}(B)).$$

For a permutation $\pi : \mathbb{N} \rightarrow \mathbb{N}$ we can define another random variable $X_\pi = (X_{\pi(1)}, X_{\pi(2)}, \dots)$. The sequence $\{X_n\}_{n \geq 0}$ is exchangeable if and only if for every π that fixes all but finitely many $i \in \mathbb{N}$, $\mu_{X_\pi} = \mu_X$.

For example, if $\{X_n\}_{n \geq 0}$ is an exchangeable sequence, then

$$\mathbb{P}(X_2 + X_3 + X_5^2 > 20 \text{ and } X_2 < 5) = \mathbb{P}(X_1 + X_3 + X_{10000}^2 > 20 \text{ and } X_1 < 5).$$

De Finetti's theorem is a theorem about exchangeable sequences. A special case of de Finetti's theorem concerns $\{0, 1\}$ -valued exchangeable sequences.

Theorem 2 (de Finetti's theorem for coin tosses.). *Let $\{X_n\}_{n \geq 0}$ be an exchangeable sequence of $\{0, 1\}$ valued random variables. Then there exists a probability measure μ on $[0, 1]$ such that for every n and every choice of $x_1, \dots, x_n \in \{0, 1\}$,*

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \int_0^1 \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i} \mu(d\theta).$$

Furthermore, the limit $\Theta = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i$ exists almost surely and the random variable Θ is distributed according to μ .

Informally, conditional on the random variable Θ , X_1, \dots, X_n are i.i.d. random variables with $\mathbb{P}(X_i = 1 | \Theta = \theta) = \theta$.

3.1 The exchangeable σ -algebra

Let X_1, X_2, \dots be any sequence of random variables. Recall that the σ -algebra $\mathcal{G} = \sigma(X_1, X_2, \dots)$ consists of all set of the form $X^{-1}(B)$ where $X = (X_1, X_2, \dots)$ and B is a measurable subset of $\mathbb{R}^{\mathbb{N}}$.

Definition 2. A set $B \subseteq \mathbb{R}^{\mathbb{N}}$ is *invariant under permutations of the first n coordinates* if for all permutations π of $\{1, \dots, n\}$, for all $x = (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}}$, if $x \in B$, then $(x_{\pi(1)}, \dots, x_{\pi(n)}, x_{n+1}, \dots) \in B$.

For example $\{x_1 + x_2 + x_3 > 5\}$ is invariant under permutations of the first 3 coordinates. The set $\{x_1 + x_2 + x_3^2 > 5\}$ is invariant under permutations of the first 2 coordinates but not invariant under permutations of the first 2 coordinates.

Definition 3. Let $\{X_n\}_{n \geq 0}$ be a sequence of random variables and let $\mathcal{G} = \sigma(X_1, X_2, \dots)$. Define \mathcal{E}_n be the σ -algebra of all $A \in \mathcal{G}$ such that $A = X^{-1}(B)$ for some $B \subseteq \mathbb{R}^{\mathbb{N}}$ that is invariant under permutations of the first n coordinates. Also define $\mathcal{E} = \bigcap_{n=1}^{\infty} \mathcal{E}_n$. The σ -algebra \mathcal{E} is called the *exchangeable σ -algebra* of $\{X_n\}_{n \geq 0}$.

We are now ready to state the general version of de Finetti's theorem.

Theorem 3 (de Finetti's theorem). *Let $\{X_n\}_{n \geq 0}$ be an exchangeable sequence of random variables and let \mathcal{E} be the exchangeable σ -algebra of $\{X_n\}_{n \geq 0}$. Then $\{X_n\}_{n \geq 0}$ are i.i.d. given \mathcal{E} meaning that for every n and every choice of Borel sets A_1, \dots, A_n , we have*

$$\begin{aligned} \mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n | \mathcal{E}) &= \prod_{i=1}^n \mathbb{P}(X_i \in A_i | \mathcal{E}) \\ &= \prod_{i=1}^n \mathbb{P}(X_1 \in A_i | \mathcal{E}). \end{aligned}$$