MIT 18.675 FALL 2019. LECTURE TOPICS

The notes below give a brief (not guaranteed to be exhaustive) summary of the main topics covered in each lecture, and relevant reading. It will be updated throughout the semester.

CONTENTS

1. (09/04) Introduction to measure theory 2
2. (09/09) Lebesgue–Stieltjes measures on the real line 2
3. (09/11) Lebesgue–Stieltjes measures on the real line, continued 2
4. (09/16) Random variables (measurable functions) and expectation (Lebesgue integral) 3
5. (09/18) Integral convergence theorems, change of variables formula 4
6. (09/23) Product measures; Tonelli and Fubini theorems 4
7. (09/25) Measures on infinite product spaces 4
8. (09/30) Independence; basic moment inequalities; $L^2$ weak law of large numbers 5
9. (10/02) Laws of large numbers 6
References 7
1. (09/04) Introduction to measure theory

1. Example: a game with an infinite sequence of boxes, axiom of choice, and probabilities. (This specific example was told to me by Persi Diaconis, and I don’t know the original source. You can see discussion online at https://mathoverflow.net/questions/151286/probabilities-in-a-riddle-involving-axiom-of-choice.)

2. Vitali’s construction, which shows that there exists no \( \mu : \mathcal{P}([0,1]) \to [0,1] \) satisfying (i) \( \mu([a,b]) = b - a \), (ii) \( \mu \) is countably additive, (iii) \( \mu \) is translation invariant. See [Dur19, §A.2] for a similar example.

3. Formal definition of a measure space \((\Omega, \mathcal{F}, \mu)\) (state space, \(\sigma\)-field or \(\sigma\)-algebra, measure). Some consequences of the definition: \(\{\emptyset, \Omega\} \in \mathcal{F}, \mu(\emptyset) = 0\); \(\mu\) is monotone (if \(A, B \in \mathcal{F}\) with \(A \subseteq B\) then \(\mu(A) \leq \mu(B)\)); continuity from below; continuity from above. Moreover \(\mu\) is **countably subadditive** over \(\mathcal{F}\): if \(A, A_i \in \mathcal{F}\) and \(A\) is contained in the countable union of the \(A_i\), then

\[
\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i) .
\]

Check for yourself that you can prove all these properties!

4. The **Borel \(\sigma\)-field over \(\mathbb{R}\)**, denoted \(\mathcal{B}_\mathbb{R}\), is the smallest \(\sigma\)-field over \(\mathbb{R}\) that contains the open intervals,

\[
\mathcal{I} = \{(a, b) : -\infty < a < b < \infty\}.
\]

We say that \(\mathcal{B}_\mathbb{R}\) is the \(\sigma\)-field generated by \(\mathcal{I}\), denoted \(\mathcal{B}_\mathbb{R} = \sigma(\mathcal{I})\). Check for yourself that it is equivalent to \(\mathcal{B}_\mathbb{R} = \sigma(\mathcal{P})\) where

\[
\mathcal{P} = \{(a, b] \cap \mathbb{R} : -\infty < a < b \leq \infty\} .
\]

(Note that \(\mathcal{I}\) contains unbounded intervals of the form \((-\infty, b]\) and \((a, \infty)\).) Check that it is also equivalent to \(\mathcal{B}_\mathbb{R} = \sigma(\mathcal{F})\) where \(\mathcal{F}\) is the collection of all open sets in the standard topology on \(\mathbb{R}\).

Reading: [Dur19, §A.2], and first part of [Dur19, §1.1].

2. (09/09) Lebesgue–Stieltjes measures on the real line

This lecture was given by Prof. Subhrabata Sen.

1. A **Stieltjes measure function** on \(\mathbb{R}\) is a function \(F : \mathbb{R} \to \mathbb{R}\) which is nondecreasing and right-continuous.

**Theorem 1.** For any Stieltjes measure function \(F\) on \(\mathbb{R}\), there is a unique measure \(\mu = \mu_F\) on \((\mathbb{R}, \mathcal{B}_\mathbb{R})\) satisfying

\[
\mu((a, b] \cap \mathbb{R}) = F(b) - F(a)
\]

for all \(-\infty \leq a \leq b \leq \infty\), where the values of \(F(\infty)\) and \(F(-\infty)\) are defined by continuity. In the case \(F(x) = x\), the corresponding \(\mu\) is called **Lebesgue measure** on \((\mathbb{R}, \mathcal{B}_\mathbb{R})\).

2. Note that condition (1) defines \(\mu : \mathcal{I} \to [0, \infty]\). Let

\[
\mathcal{A} = \left\{ A \subseteq \mathbb{R} : A \text{ is a disjoint union of finitely many elements of } \mathcal{I} \right\} .
\]

This is an **algebra** (closed under complementation and finite union); it is the smallest algebra containing \(\mathcal{I}\).

**Part I** of the proof of Theorem 1: there is a unique \(\mu : \mathcal{A} \to [0, \infty]\) which extends \(\mu : \mathcal{I} \to [0, \infty]\), and is countably additive over \(\mathcal{A}\).

Reading: [Dur19, §1.1].

3. (09/11) Lebesgue–Stieltjes measures on the real line, continued

This lecture was given by Prof. Subhrabata Sen.

1. **Part II** of the proof of Theorem 1: there is a unique \(\mu : \mathcal{B}_\mathbb{R} \to [0, \infty]\) which extends \(\mu : \mathcal{A} \to [0, \infty]\) and is a measure on \(\mathcal{B}_\mathbb{R}\) (i.e., is countably additive over \(\mathcal{B}_\mathbb{R}\)). In the case that \(\mu\) is a finite measure (\(\mu(\Omega) < \infty\)), this is a special case of a more general theorem:

**Theorem 2** (Carathéodory extension theorem). *Suppose \(\mathcal{A}\) is an algebra over \(\Omega\), and \(\mu : \mathcal{A} \to [0, \infty]\) (in particular, \(\mu(\Omega) < \infty\)) is countably additive over \(\mathcal{A}\). Then there is a unique \(\mu : \sigma(\mathcal{A}) \to [0, \infty]\) which extends \(\mu : \mathcal{A} \to [0, \infty]\) and is a measure on \(\sigma(\mathcal{A})\) (i.e., is countably additive over \(\sigma(\mathcal{A})\)).

2. Proof of uniqueness in Theorem 2: this is based on **Dynkin’s \(\pi\)-\(\lambda\) theorem**.
3. Proof of existence in Theorem 2: this is based on the construction of outer measure. In the particular setting of Theorem 1 (which is less general than the setting of Theorem 2), the outer measure can be defined as

$$\mu^*(A) = \inf \left\{ \sum_{i=1}^{\infty} \mu(E_i) : E_i = (a_i, b_i] \text{ and } A \subseteq \bigcup_{i=1}^{\infty} E_i \right\},$$

where $$\mu(E_i) = F(b_i) - F(a_i)$$. Then let

$$\mathcal{F}_F \equiv \left\{ A \subseteq \Omega : \mu^*(S) = \mu^*(S \cap A) + \mu^*(S \setminus A) \text{ for all } S \subseteq \Omega \right\}.$$

We showed in class that $$\mathcal{F}_F$$ is a σ-algebra, and the restriction of $$\mu^*$$ to $$\mathcal{F}_F$$ is a measure $$\mu = \mu_F$$. In the case $$F(x) = x$$, $$\mathcal{F}_F$$ is the Lebesgue σ-algebra, commonly denoted $$\mathcal{L}_R$$; and $$\mu_F$$ is the Lebesgue measure, commonly denoted $$\lambda$$ or Leb. Note $$\mathcal{L}_R \supseteq \mathcal{B}_R$$, so $$\lambda$$ further restricts to a measure on $$(\mathbb{R}, \mathcal{A})$$ which is also called Lebesgue measure.

4. The precise relation between $$\mathcal{L}_R$$ and $$\mathcal{B}_R$$ is as follows: any $$A \in \mathcal{L}_R$$ can be expressed as $$A = B \cup N$$ where $$B \in \mathcal{B}$$ and $$N \subseteq B' \in \mathcal{B}$$ with $$\lambda(B') = 0$$. This statement is a consequence of [Dur19, Theorem A.2.2]. We did not cover it in lecture (although we may see it in a future lecture or homework).

Reading: [Dur19, §A.1].

4. (09/16) Random variables (measurable functions) and expectation (Lebesgue integral)

1. Let $$(\Omega, \mathcal{F})$$ be a measurable space. We say $$f$$ is a simple function on $$(\Omega, \mathcal{F})$$ if $$f : \Omega \to \mathbb{R}$$ can be expressed as

$$f = \sum_{i=1}^{n} c_i 1_{A_i},$$

(2)

for $$c_i \in \mathbb{R}$$ and $$A_i \in \mathcal{F}$$. A measurable function on $$(\Omega, \mathcal{F})$$ is any pointwise limit of simple functions: $$f = \lim_n f_n$$ where $$f_n$$ are simple functions on $$(\Omega, \mathcal{F})$$. A measurable mapping from $$(\Omega, \mathcal{F})$$ to $$(S, \mathcal{G})$$ is a map $$g : \Omega \to S$$ such that for all $$B \in \mathcal{G}$$, the preimage $$g^{-1}(B)$$ belongs to $$\mathcal{F}$$. To emphasize measurability of $$g$$, we often write $$g : (\Omega, \mathcal{F}) \to (S, \mathcal{G})$$. On Homework 1: $$f$$ is a measurable function on $$(\Omega, \mathcal{F})$$ if and only if it is a measurable mapping from $$(\Omega, \mathcal{F})$$ to $$(\mathbb{R}, \mathcal{B}_R)$$.

2. If $$\mu$$ is any measure on $$(\Omega, \mathcal{F})$$ and $$g : (\Omega, \mathcal{F}) \to (S, \mathcal{G})$$, we obtain a pushforward measure $$\nu$$ on $$(S, \mathcal{G})$$,

$$\nu(B) = \mu(g^{-1}(B)) = \mu\left( \{ \omega \in \Omega : g(\omega) \in B \} \right).$$

Some standard notations: $$\nu = g_* \mu = g \circ \mu^{-1}$$.

3. Let $$(\Omega, \mathcal{F})$$ be a σ-finite measure space. If $$f$$ is a simple function on $$(\Omega, \mathcal{F})$$ as in (2), with the additional condition that $$\mu(A_i) < \infty$$ for all $$i$$, then we define its Lebesgue integral

$$\int_{\Omega} f \, d\mu = \sum_{i=1}^{n} c_i \mu(A_i).$$

If $$f$$ is a nonnegative measurable function on $$(\Omega, \mathcal{F})$$, we define its Lebesgue integral

$$\int_{\Omega} f \, d\mu = \sup \left\{ \int_{\Omega} h \, d\mu : 0 \leq h \leq f, \ h \text{ simple function with } \mu(h) < \infty \right\},$$

Finally, if $$f$$ is a general measurable function on $$(\Omega, \mathcal{F})$$, we define its Lebesgue integral

$$\int_{\Omega} f \, d\mu = \int_{\Omega} f^+ \, d\mu - \int_{\Omega} f^- \, d\mu$$

with the caveat that $$\infty - \infty = 0$$. This completes the general definition of the Lebesgue integral

$$\int_{\Omega} f \, d\mu = \int_{\Omega} f(\omega) \, d\mu(\omega).$$

It can be finite or $$\pm \infty$$. Note from the definition that the Lebesgue integral of $$f$$ is finite if and only if

$$\int_{\Omega} |f| \, d\mu < \infty,$$
and in this case we call $f$ integrable. Basic properties of the integral (monotonicity, linearity, etc.): see the results Lemma 1.4.3, Lemma 1.4.5, Theorem 1.4.7 in [Dur19].

4. A random variable is a measurable function $X$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Its distribution or law is $\mathcal{L}_X = X_*\mathbb{P}$. A useful notation: for $B \in \mathcal{B}_\mathbb{R}$,

$$X^{-1}(B) \equiv \{ \omega \in \Omega : X(\omega) \in B \} \equiv \{ X \in B \}.$$  

This allows us to write $\mathcal{L}_X(B)$ in a more intuitively natural way as simply $\mathbb{P}(X \in B)$. The expectation or mean of $X$ is its Lebesgue integral,

$$EX = \int_\Omega X d\mathbb{P} = \int_\Omega X(\omega) d\mathbb{P}(\omega).$$

Note from the definitions that $EX$ is finite if and only if $\mathbb{E}|X|$ is finite.

Reading: [Dur19, §1.2–1.4]. The Lebesgue integral subsumes the Riemann integral: see for instance the statement of [Tao11, Exercise 1.3.17] (the book gives plenty of hints to prove this statement, which we will take for granted in this class).

5. (09/18) Integral convergence theorems, change of variables formula

1. Modes of convergence: $f_n \rightarrow f$ pointwise, $\mu$-almost everywhere ($\mu$-a.e.), in $\mu$-measure.
2. Integral convergence theorems: bounded convergence theorem, Fatou’s lemma, monotone convergence theorem, dominated convergence theorem.
3. Application of the monotone convergence theorem to a change of variables formula: suppose

$$Y = f(X) : (\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{X} (S, \mathcal{G}, \mu) \overset{f}{\rightarrow} (\mathbb{R}, \mathcal{B}, \nu)$$

where $\mu = X_*\mathbb{P}$ and $\nu = f_*\mu = Y_*\mathbb{P}$. Provided either $f \geq 0$ or $Y$ is integrable, we have $I_\Omega = I_S$ where

$$I_\Omega \equiv \mathbb{E}Y \equiv \int_\Omega Y d\mathbb{P} \equiv \int_\Omega f(X(\omega)) d\mathbb{P}(\omega),$$

$$I_S \equiv \int_S f d\mu \equiv \int_S f(x) d(\mathbb{P} \circ f^{-1})(x)$$

Proof uses a standard technique: start with indicators, then extend to simple functions, then nonnegative functions (using the monotone convergence theorem), then general functions.

Reading: [Dur19, §1.5–1.6].

6. (09/23) Product measures; Tonelli and Fubini theorems

1. Given two $\sigma$-finite measure spaces $(S, \mathcal{G}, \lambda)$ and $(T, \mathcal{H}, \rho)$, we defined the product measure space $(\Omega, \mathcal{F}, \mu) = (S \times T, \mathcal{G} \otimes \mathcal{H}, \lambda \otimes \rho)$.
2. Tonelli and Fubini theorems: for measurable $f : \Omega \rightarrow \mathbb{R}$, provided each $f \geq 0$ or $\int |f| d\mu < \infty$, we have

$$\int_S \int_T f(x, y) d\rho(y) d\lambda(x) = \int_S \int_T f(x, y) d\mu(x, y) = \int_T \int_S f(x, y) d\lambda(x) d\rho(y).$$

(It is part of the content of the theorem that the left-hand side and right-hand side are well-defined quantities.)

Reading: [Dur19, §1.7].

7. (09/25) Measures on infinite product spaces

1. Let $(\Omega_\alpha, \mathcal{F}_\alpha)$ be measurable spaces indexed by $\alpha \in I$ (index set, possibly uncountable). Their product:

$$(\Omega, \mathcal{F}) = \left( \prod_{\alpha \in I} \Omega_\alpha, \bigotimes_{\alpha \in I} \mathcal{F}_\alpha \right)$$

where $\mathcal{F}$ is the minimal $\sigma$-field such that the coordinate projections $\pi_\alpha : \Omega \rightarrow \Omega_\alpha$ are measurable. For $J \subseteq I$ denote the partial products

$$\Omega_J \equiv \prod_{\alpha \in J} \Omega_\alpha, \quad \mathcal{F}_J \equiv \bigotimes_{\alpha \in J} \mathcal{F}_\alpha.$$
2. If $\mathcal{P}$ is a probability measure on this $(\Omega, \mathcal{F})$, its **finite-dimensional marginals** are the measures $(\pi_j)_n \mathcal{P}$. This means that the marginal probability of $E_j \in \mathcal{F}_j$ is

$$\left((\pi_j)_n \mathcal{P}\right)(E_j) = \mathcal{P}\left((\pi_j)^{-1}(E_j)\right) = \mathcal{P}(E_j \times \Omega_{J \setminus j}).$$

They have to be **consistent**: if $J' \subseteq J \subseteq I$ and $\pi_{J \setminus J'}$ is the projection from $\Omega_J$ to $\Omega_{J'}$, then

$$(\pi_{J \setminus J'})_n \mathcal{P} = (\pi_{J \setminus J'})_n \left((\pi_j)_n \mathcal{P}\right).$$

Both the theorems from this lecture go in reverse: given a consistent family of finite-dimensional marginals $\{\mathcal{P}_j : \text{finite } J \subseteq I\}$, they construct a measure $\mathcal{P}$ on $(\Omega, \mathcal{F})$ with these finite-dimensional marginals. Note: $\{\mathcal{P}_j : \text{finite } J \subseteq I\}$ is often also called the **finite-dimensional distributions**, abbreviated f.d.d.

3. A useful variant of the criterion for the Carathéodory extension theorem (Theorem 2 above):

**Lemma 3.** Suppose $\mathcal{A}$ is an algebra of sets over $\Omega$, and that $\mu : \mathcal{A} \rightarrow [0, \infty)$ is finitely additive over $\mathcal{A}$. If for any sequence $B_n \in \mathcal{A}$ with $B_n \downarrow \emptyset$ we have $\mu(B_n) \downarrow \emptyset$, then $\mu$ is countably additive over $\mathcal{A}$.

Specialization of Lemma 3 to the infinite product setting (display (3) above): it is enough to prove the criterion in the scenario that we have $Q = \{1, 2, \ldots, \} \subseteq I$ countable and $B_n \downarrow \emptyset$ of the form

$$B_n = \tilde{B}_n \times \prod_{i=n+1}^{\infty} \Omega_i \times \prod_{\alpha \in I \setminus \Omega} \Omega_{\alpha} = \tilde{B}_n \times \Omega_{\Omega \setminus \{n\}} \times \Omega_{\text{rest}}$$

where $[n] \equiv \{1, \ldots, n\}$ and $\tilde{B}_n \in \mathcal{F}_{[n]}$.

4. Product measures on an infinite-dimensional spaces:

**Theorem 4** (Ionescu–Tulcea theorem). Let $(\Omega_{\alpha}, \mathcal{F}_{\alpha}, \mathcal{P}_{\alpha})$ be probability spaces indexed by $\alpha \in I$. There is a unique probability measure $\mathcal{P}$ on the product space $(\Omega, \mathcal{F})$ (as in (3)) with finite-dimensional marginals $(\pi_j)_n \mathcal{P} = \bigotimes_{\alpha \in J} \mathcal{P}_{\alpha}$.

5. Non-product measures on an infinite-dimensional spaces:

**Theorem 5** (Kolmogorov extension theorem). Let $(\Omega_{\alpha}, \mathcal{F}_{\alpha})$ be metric spaces with the Borel $\sigma$-field. Suppose $\{\mathcal{P}_j : \text{finite } J \subseteq I\}$ is a consistent family of finite-dimensional distributions, and that each $\mathcal{P}_J$ is an inner regular measure on $(\Omega_J, \mathcal{F})$. Then there is a unique probability measure $\mathcal{P}$ on $(\Omega, \mathcal{F})$ with finite-dimensional marginals $(\pi_j)_n \mathcal{P} = \mathcal{P}_J$.

Reading: [Dur19, Thm. 2.1.14 and §A.3], and [Kal02, Ch. 5] (especially Thm. 5.14, Thm. 5.17 and Cor. 5.18).

8. (09/30) Independence; basic moment inequalities; $L^2$ weak law of large numbers

1. Independence (of events, of collections of events, and of random variables). Connection with previous week:

   a. If the probability space comes with a product structure

   $$\left.\left(\prod_{\alpha \in I} \Omega_{\alpha}, \bigotimes_{\alpha \in I} \mathcal{F}_{\alpha} \right) \bigotimes \mathcal{P}_{\alpha}\right),$$

   and $X_{\alpha}(\omega) = f_\alpha(\omega_{\alpha})$ where $f_\alpha$ is a measurable function on $\Omega_{\alpha}$, then $(X_{\alpha} : \alpha \in I)$ is a collection of independent random variables on $(\Omega, \mathcal{F}, \mathcal{P})$.

   b. If $(\Omega, \mathcal{F}, \mathcal{P})$ is a general probability space (i.e., not necessarily equipped with an explicit product structure) and we are told that $(X_{\alpha} : \alpha \in I)$ is a collection of random variables on $(\Omega, \mathcal{F}, \mathcal{P})$, we can define $X(\omega) \equiv (X_\alpha(\omega) : \alpha \in I)$. This gives a mapping

   $$\begin{align*}
   (\Omega, \mathcal{F}, \mathcal{P}) \rightarrow & \ X \left(\prod_{\alpha \in I} \mathbb{R}, \bigotimes_{\alpha \in I} \mathcal{B}_{\mathbb{R}} \right) \ 
   \mathcal{L}_X \equiv X_{\mathbb{P}} = \mathcal{P} \circ X^{-1}
   \end{align*}$$

   where $\mathcal{L}_X \equiv X_{\mathbb{P}} = \mathcal{P} \circ X^{-1}$ is the law of $X$ (as defined previously). Then $(X_{\alpha} : \alpha \in I)$ is a collection of independent random variables if and only if $\mathcal{L}_X$ is a product measure.
2. Moments of random variables: for $p \in (0, \infty)$ and $X$ a random variable we define

$$\|X\|_p \equiv \|X\|_{L^p(\Omega, \mathcal{F}, \mathbb{P})} \equiv \left\{ \int |X(\omega)|^p \, d\mathbb{P}(\omega) \right\}^{1/p} \equiv \mathbb{E}(|X|^p)^{1/p}.$$ 

We write $L^p \equiv L^p(\Omega, \mathcal{F}, \mathbb{P})$ for the collection of $X$ with $\|X\|_p < \infty$. $L^p$ monotonicity: if $r \leq p$ then $\|X\|_r \leq \|X\|_p$, so $L^p(\Omega, \mathcal{F}, \mathbb{P}) \subseteq L^r(\Omega, \mathcal{F}, \mathbb{P})$. Be care that for $L^p$ monotonicity it is essential to be on a probability space. The $\ell_p$ sequence spaces are nested in the opposite direction: if $r \leq p$ then

$$\|x\|_r \equiv \left( \sum_i |x_i|^r \right)^{1/r} \geq \left( \sum_i |x_i|^p \right)^{1/p} = \|x\|_p,$$

and so the $\ell_p$ unit ball contains the $\ell_r$ unit ball. Finally, the classical $L^p$ function spaces are not nested at all: if

$$\|f\|_p \equiv \|f\|_{L^p(\mathbb{R})} \equiv \left\{ \int |f(x)|^p \, dx \right\}^{1/p}$$

and $L^p(\mathbb{R}) \equiv \{ f : \|f\|_p < \infty \}$, both $L^p(\mathbb{R}) \setminus L^r(\mathbb{R})$ and $L^r(\mathbb{R}) \setminus L^p(\mathbb{R})$ are nonempty for $r \neq p$.

3. If $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, we define its variance $\text{Var} X = \mathbb{E}[(X - \mathbb{E}X)^2]$. Cauchy–Schwarz inequality: if $X, Y \in L^2$ then $\|XY\|_1 = \mathbb{E}|XY| \leq \|X\|_2 \|Y\|_2$. Consequently, for $X, Y \in L^2$ we can define their covariance

$$\text{Cov}(X, Y) \equiv \mathbb{E}\left((X - \mathbb{E}X)(Y - \mathbb{E}Y)\right), \quad |\text{Cov}(X, Y)| \leq \left\{ \text{Var}(X)(\text{Var} Y) \right\}^{1/2}.$$

We say $X$ and $Y$ are uncorrelated if $\text{Cov}(X, Y) = 0$.

4. Markov inequality; Chebychev inequality; and a simple version of the $L^2$ weak law of large numbers: if $X, X_i$ are pairwise uncorrelated and identically distributed, and $S_n \equiv X_1 + \ldots + X_n$, then $S_n/n \to \mathbb{E}X$ in $L^2$, hence also in probability (by Chebychev). We noted that the $L^2$ convergence is a restatement of a simple geometric fact: if $(v_1, \ldots, v_n)$ are orthonormal vectors, then their average has small euclidean norm:

$$\left\| \frac{1}{n} \sum_{i=1}^n v_i \right\|_2 \leq \frac{1}{n^{1/2}}$$

($v_i$ corresponds to $X_i - \mathbb{E}X$).

Reading: PTE 2.1.

9. (10/02) Laws of large numbers

Setting for entirety of this lecture: triangular array with $n$-th row given by $(X_{n,k} : 1 \leq k \leq n)$. Assume independence within each row. Important special case is $X_{n,k} = X_k$ where $X_k$ are i.i.d. Sum of $n$-th row is

$$S_n \equiv \frac{1}{n} \sum_{k=1}^n X_{n,k}.$$

Interested in convergence of $S_n/b_n$ or $(S_n - \mathbb{E}S_n)/b_n$ for $b_n$ deterministic. WLLN stands for “weak law of large numbers.” SLLN stands for “strong law of large numbers.”

1. $L^2$ WLLN for triangular arrays. If $\text{Var} X_{n,k} \leq C$ uniformly, then $(S_n - \mathbb{E}S_n)/b_n \to 0$ in $L^2$ as $n \to \infty$, and hence also in probability (by Chebychev).

2. WLLN for triangular arrays. Suppose $b_n \to \infty$ such that

(i) $\lim_{n \to \infty} \sum_{k=1}^n \mathbb{P}(\{X_{n,k} > b_n\}) = 0,$

(ii) $\lim_{n \to \infty} \sum_{k=1}^n \frac{\mathbb{E}((X_{n,k})^2)}{(b_n)^2} = 0.$

Let $Y_{n,k} \equiv X_{n,k}1[|X_{n,k}| \leq b_n]$ and define the truncated row sums $T_n \equiv Y_{n,1} + \ldots + Y_{n,n}$. Then

$$(S_n - \mathbb{E}T_n)/b_n \to 0$$

in probability. Proof synopsis: $\mathbb{P}(S_n \neq T_n) \to 0$ by (i), $(T_n - \mathbb{E}T_n)/b_n \to 0$ in $L^2$ by (ii), hence also in probability by Chebychev.
3. **WLLN for i.i.d. sequences.** Let $X, X_k$ i.i.d. with $\lim_{x \to \infty} x P( |X| \geq x ) = 0$. Define $\mu_n \equiv E(X; |X| \leq n)$. Then $(S_n/n - \mu_n) \to 0$ in probability. Note that $E X$ need not be defined. The proof is an application of the previous result together with the dominated convergence theorem.

4. **SLLN for i.i.d. sequences.** Let $X, X_k$ i.i.d. with $E X = \mu$ finite, then $S_n/n \to \mu$ almost surely.

Reading: PTE 2.2–2.4.

Tentative plan for next week (October 7 and 9): introduction to large deviations theory.

### References

