MATH 275C

Lecture 1: March 31, 2025

- Kolmogorov's model of probability inside measure theory
- Integration, expectation, variance, moments
- Independence and product measures
- Basic convergence theorems: WLLN, SLLN, random series
- Weak convergence of prob. measures/random variables
- Characteristic function and Central Limit Theorem

- Quantitative CLTs: Lindeberg and Stein methods
- Stable convergence and stable laws
- Infinite divisibility & Lévy-Khinchin formula
- Conditional expectation and probability
- Uniform integrability
- Martingales: convergence, optional stopping, etc
- Exchangeability & de Finetti theorem
- von Neumann/Birkhoff's Ergodic Theorems, ergodicity

- Wrapping up ergodic theory started in 275B
- Discrete-time Markov chains, random walks
- Continuous-time processes: Renewals, Markov chains
- Processes with independent increments: Brownian motion
- General Markov processes

Motivated by Boltzmann's Ergodic Hypothesis, we introduced:

Definition: Let $(\mathscr{X}, \mathcal{G}, \mu)$ be a measure space. A map $\varphi\colon \mathscr{X} \to \mathscr{X}$ is a **measure preserving transformation** (m.p.t.) if

- φ is measurable (i.e., $\varphi^{-1}(\mathcal{G}) \subseteq \mathcal{G}$)
- φ preserves μ (i.e., $\mu \circ \varphi^{-1} = \mu$)

We call $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ a **measure preserving system** (m.p.s.)

• **Hamiltonian flow**: Let $t \mapsto (q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t))$ solve the ODEs

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i}$$

where $H \in C^2(\mathbb{R}^{2n})$ is the **Hamiltonian**

Liouville's theorem: Let $t > 0$. Then $\varphi\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ defined by

$$\varphi(q(0), p(0)) := (q(t), p(t))$$

preserves the Lebesgue measure on $\mathbb{R}^{2n}$

Note: The flow leaves H constant so even the "surface measure" on $\{H = E\}$ preserved for each E

• **Stationary sequences**: Let $(S,\Sigma) :=$ measurable space and $\{X_k\}_{k\geqslant 0} := S$-valued random variables on a probability space $(\Omega,\mathcal{F},P)$ such that

$$\forall n,k \geqslant 0\colon\ (X_n,\dots,X_{n+k}) \overset{\text{law}}{=} (X_0,\dots,X_k)$$

Denoting $X := \{X_k\}_{k\geqslant 0}$, the map $X\colon \Omega \to S^{\mathbb{N}}$ pushes P to a probability measure μ on $(S^{\mathbb{N}},\Sigma^{\otimes\mathbb{N}})$. The **left shift** $\theta\colon S^{\mathbb{N}} \to S^{\mathbb{N}}$ defined by

$$\theta\big(\{x_k\}_{k\geqslant 0}\big) = \{x_{k+1}\}_{k\geqslant 0}$$

is an m.p.t. on $(S^{\mathbb{N}},\Sigma^{\otimes\mathbb{N}},\mu)$

Examples: i.i.d., stationary Markov chains, etc

- **Rotation of a circle**: $\mathscr{X} := [0,1)$, $\mathcal{G} := \mathcal{B}([0,1))$. Then

$$\varphi(x) := x + \alpha \bmod 1$$

preserves $\mu :=$ Lebesgue measure for each $\alpha \in \mathbb{R}$

- **Continued fractions**: $\mathscr{X} := [0,1] \setminus \mathbb{Q}$, $\mathcal{G} := \mathcal{B}(\mathscr{X})$. Then

$$\varphi(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

preserves $\mu(\mathrm{d}x) := \frac{1}{1+x}\mathrm{d}x$. Name stems from representation

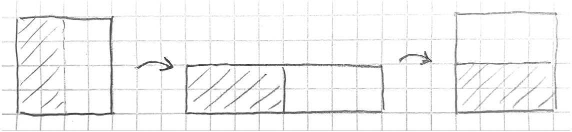
$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \dots}}}}$$

on which φ acts as $(a_1, a_2, \dots) \mapsto (a_2, a_3, \dots)$

• **Baker's transform**: $\mathscr{X} := [0,1) \times [0,1)$, $\mathcal{G} = \mathcal{B}(\mathscr{X})$. Then

$$\varphi(x,y) = \begin{cases} (2x, \frac{1}{2}y), & \text{if } 0 \leqslant x < 1/2 \\ (2x-1, \frac{1}{2}(y+1)), & \text{if } 1/2 \leqslant x < 1 \end{cases}$$

preserves μ:= Lebesgue measure.



Coded by Bernoulli's:

$$x = \sum_{n \geqslant 0} \frac{X_n}{2^{n+1}} \quad \text{and} \quad y = \sum_{n<0} \frac{X_n}{2^{-n}}$$

Then φ acts as a left shift on $\{X_n\}_{n\in\mathbb{Z}}$

Note: φ = m.p.t. $\Rightarrow Tf(x) := f \circ \varphi$ obeys $\|Tf\|_p = \|f\|_p$

Definition: A linear map $T\colon \mathcal{B} \to \mathcal{B}$ on Banach space $\mathcal{B}$ is a **contraction** if $\|Tf\| \leqslant \|f\|$ for all $f \in \mathcal{B}$

Theorem (Mean Ergodic Theorem)

Let T be contraction on Hilbert space $\mathcal{H}$. Then

$$\forall f \in \mathcal{H}\colon \quad \frac{1}{n}\sum_{k=0}^{n-1} T^k f \underset{n\to\infty}{\longrightarrow} Pf$$

where $P :=$ orthogonal projection on $\operatorname{Ker}(1-T)$

Idea: Easy to check for $f := g + h - Th$ with $Tg = g$. Then need

$$\{g + h - Th\colon h, g \in \mathcal{H},\, Tg = g\} \text{ dense in } \mathcal{H}$$

Note: Extends to all **reflexive** Banach spaces

Theorem (Pointwise Ergodic Theorem)

Let $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$:= *m.p.s. with* $\mu(\mathscr{X}) < \infty$. *Then* $\forall f \in L^1\ \exists \bar{f} \in L^1$:

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ \varphi^k \underset{n\to\infty}{\longrightarrow} \bar{f} \qquad \mu\text{-a.e. \& in } L^1$$

Moreover, $\bar{f} \circ \varphi = \bar{f}$ *μ-a.e. and* $\|\bar{f}\|_1 \leqslant \|f\|_1$

"Proof of the Ergodic Theorem" by G.D. Birkhoff, PNAS 1931

Theorem (Maximal Ergodic Theorem)

Let $(\mathscr{X}, \mathcal{G}, \mu)$:= *measure space and let* $T\colon L^1 \to L^1$ *be a positivity-preserving contraction. For* $f \in L^1$, *denote*

$$f^\star := \sup_{n\geqslant 1} \frac{1}{n} \sum_{k=0}^{n-1} T^k f$$

Then

$$\int_{\{f^\star>0\}} f \mathrm{d}\mu \geqslant 0$$

Wiener 1939, Yoshida & Kakutani 1939
A slick proof by Garsia 1965
We will give a different argument later

Corollary

Suppose T as above and also $\mu(\mathscr{X}) < \infty$. Then

$$\left\{ f \in L^1 \colon \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} T^k f \text{ exists a.e.} \right\} \text{ is closed in } L^1$$

Proof: For $h \in L^1$ abbreviate $A_n h := \frac{1}{n}\sum_{k=0}^{n-1} T^k h$ and observe $|A_n h| \leqslant A_n |h|$. This implies

$$\{\limsup_{n\to\infty} |A_n h| > \epsilon\} \subseteq \{|h|^\star > \epsilon\}$$

and $(|h| - \epsilon)^\star = |h|^\star - \epsilon$. Maximal Ergodic Theorem gives

$$\int_{|h|^\star - \epsilon} \big(|h| - \epsilon\big) \mathrm{d}\mu \geqslant 0$$

Need $\mu(X) < \infty$ here!

implying

$$\mu\big(|h|^\star > \epsilon\big) \leqslant \frac{1}{\epsilon} \|h\|_1$$

This is a maximal inequality as in Hardy-Littlewood, Doob, etc

Now consider $\{f_k\}$ such that $f_k \to f$ in L^1 and $\bar{f}_k := \lim_{n\to\infty} A_n f_k$ exists a.e. for all $k \geqslant 1$. The above shows

$$\bar{f}_k - \epsilon \leqslant \liminf_{n\to\infty} A_n f \leqslant \limsup_{n\to\infty} A_n f \leqslant \bar{f}_k + \epsilon \quad \text{a.e.}$$

on $\{|f_k - f|^\star \leqslant \epsilon\}$ and thus, by the maximal inequality,

$$\mu\Big(\limsup_{n\to\infty} A_n f - \liminf_{n\to\infty} A_n f > 2\epsilon\Big) \leqslant \mu\big(|f_k - f|^\star > \epsilon\big) \leqslant \frac{1}{\epsilon}\|f_k - f\|_1$$

Taking $k \to \infty$ we get that $\lim_{n\to\infty} A_n f$ exists a.e. □

Theorem (Spatial Ergodic Theorem)

Suppose $(\mathscr{X}, \mathcal{G}, \mu)$ *(with* $\mu(\mathscr{X})$ *possibly infinite) and, given* $d \geqslant 1$*, let* $\varphi_1, \ldots, \varphi_d$ *be m.p.t.'s on* $(\mathscr{X}, \mathcal{G}, \mu)$*. Assume* $\varphi_1, \ldots, \varphi_d$ *commute,*

$$\forall i, j = 1, \ldots, d\colon \;\; \varphi_i \circ \varphi_j = \varphi_j \circ \varphi_i$$

For all $f \in L^1$ *there exists* $\bar{f} \in L^1$ *such that for any sets* $\{\Lambda_n\}_{n \geqslant 1}$ *with*

$$\forall n \geqslant 1\colon \varnothing \neq \Lambda_n \subseteq [0, n)^d \cap \mathbb{Z}^d$$

and

$$\liminf_{n\to\infty} \frac{|\Lambda_n|}{n^d} > 0 \;\wedge\; \lim_{n\to\infty} \frac{|\partial \Lambda_n|}{n^d} = 0$$

we have

$$\frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} f \circ \varphi_x \underset{n\to\infty}{\longrightarrow} \bar{f} \quad \text{a.e.}$$

where $\varphi_{(x_1,\ldots,x_n)} := \varphi_1^{x_1} \circ \cdots \circ \varphi_d^{x_d}$*. If* $\mu(\mathscr{X}) < \infty$*, then the convergence is also in* L^1

Lemma (Maximal inequality in $\mathbb{Z}^d$)

Given $h\colon \mathbb{Z}^d \to \mathbb{R}$ *and* $n \geqslant 1$, *denote* $B_n(x) := x + [0,n)^d \cap \mathbb{Z}^d$. *Set*

$$h_n^\star(x) := \sup_{k=1,\dots,n} \frac{1}{|B_k(x)|} \sum_{z\in B_k(x)} \bigl|h(z)\bigr|$$

Then

$$\forall\lambda > 0\colon \quad \bigl|\{x \in B_n(0)\colon h_n^\star(x) > \lambda\}\bigr| \leqslant \frac{3^d}{\lambda} \sum_{z\in B_{2n}} \bigl|h(z)\bigr|$$

Proof: The set $\{x \in B_n(0)\colon h_n^\star(x) > \lambda\}$ is covered by

$$\Bigl\{B_k(x)\colon x \in B_n(0),\, k = 1,\dots,n,\, \sum_{x\in B_k(x)} \bigl|h(z)\bigr| > \lambda B_k(x)\Bigr\}$$

Enumerate as $B^1,\dots,B^m$ decreasingly in size with "lower-left" corner of B^i denoted as x_i. Set $\widetilde{B}^i := x_i + [-n,2n)^d \cap \mathbb{Z}^d$. Next identify indices $i_1,\dots,i_q$ as follows:

Set $i_1 := 1$. If $i_1, \dots, i_j$ defined, let

$$i_{j+1} := \min\{i > i_j \colon x_i \notin \widetilde{B}^{i_1} \cup \cdots \cup \widetilde{B}^{i_j}\}$$

if such i exists; otherwise set $q := j$ and terminate. Then

$$\big\{x \in B_n(0) \colon h_n^\star(x) > \lambda\big\} \subseteq \bigcup_{j=1}^{q} \widetilde{B}^{i_j}$$

and

$$B^{i_1}, \dots, B^{i_q} \text{ are disjoint}$$

because $B^{i_k} \cap B^{i_\ell} \neq \varnothing$ for $k < \ell$ would imply $B^{i_\ell} \subseteq \widetilde{B}^{i_k}$, in contradiction with $x_{i_\ell} \notin \widetilde{B}^{i_k}$. Now compute ...

…using the fact that $|\widetilde{B}^i| = 3^d |B^i|$ to get

$$\begin{aligned}\bigl|\{x \in B_n(0)\colon h_n^\star(x) > \lambda\}\bigr| &\leqslant \sum_{j=1}^{q} |\widetilde{B}^{i_j}| = 3^d \sum_{j=1}^{q} |B^{i_j}| \\ &\leqslant \frac{3^d}{\lambda} \sum_{j=1}^{q} \sum_{x \in B^{i_j}} \bigl|h(z)\bigr| \leqslant \frac{3^d}{\lambda} \sum_{x \in B_{2n}(0)} \bigl|h(z)\bigr|\end{aligned}$$

where we also used $B^{i_1}, \dots, B^{i_q} \subseteq B_{2n}(0)$ □

Corollary

For $f \in L^1(\mathscr{X}, \mathcal{G}, \mu)$, *set*

$$f^\star := \sup_{n \geqslant 1} \frac{1}{|B_n(0)|} \sum_{x \in B_n(0)} |f| \circ \varphi_x$$

Then

$$\forall \lambda > 0\colon \quad \mu(f^\star > \lambda) \leqslant \frac{6^d}{\lambda} \|f\|_1$$

Proof: Abbreviate $M_n f := \max_{k=1,\dots,n} \frac{1}{|B_k(0)|} \sum_{x \in B_n(0)} |f| \circ \varphi_x$.
Lemma with $h(x) := f \circ \varphi_x$ and $h_n^\star(x) = M_n f \circ \varphi_x$ gives

$$\mu(M_n f > \lambda) = \frac{1}{|B_n(0)|} \int \big|\{x \in B_n(0)\colon M_n f \circ \varphi_x > \lambda\}\big| \mathrm{d}\mu$$

$$\leqslant \frac{3^d}{\lambda} \frac{1}{|B_n(0)|} \int \sum_{x \in B_{2n}(0)} |f| \circ \varphi_x \, \mathrm{d}\mu = \frac{6^d}{\lambda} \int |f| \mathrm{d}\mu$$

Now take $n \to \infty$ □

Maximal inequality shows

$$\left\{ f \in L^1\colon \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \sum_{x\in\Lambda_n} f\circ\varphi_x \text{ exists a.e.} \right\} \text{ is closed in } L^1$$

Now use that the limit exists and is independent of $\{\Lambda_n\}$ for every element of

$$\left\{ g + \sum_{i=1}^{d} (h_i - h_i\circ\varphi_i)\colon g, h_1, \dots, h_d \in L^1 \wedge \forall i = 1,\dots,d\colon\ g = g\circ\varphi_i \right\}$$

which is dense in L^1 $\square$

TO BE CONTINUED ON WEDNESDAY . . .