PROBABILITY THEORY

MATH 154

Unit 16: Ergodic theorem

16.1. A map $T : \Omega \to \Omega$ on a probability space $(\Omega, \mathcal{A}, \mathrm{P})$ is **measurable** if $T^{-1}(A) \in \mathcal{A}$ for every $A \in \mathcal{A}$. A single random variable X defines a sequence $X_n(\omega) = X(T^n(\omega))$ of random variables where $T^n(\omega) = T(T(\dots T(\omega)))$ is the **n'th iterate**. T is called **measure preserving**, if $\mathrm{P}[T^{-1}(A)] = \mathrm{P}[A]$ for all $A \in \mathcal{A}$. It is **ergodic** if $T(A) = A$ implies $\mathrm{P}[A] = 0$ or $\mathrm{P}[A] = 1$. The map T is called **invertible**, if there exists a measurable, measure-preserving inverse T^{-1}. Then, T is called an **automorphism**. Ergodicity of T is equivalent to the statement that the linear map $U(f)(x) = f(T(x))$ has a one-dimensional eigenspace, consisting of constant functions. Given a random variable X, define $X_k = X(T^k)$ and $S_n = \sum_{k=0}^{n-1} X_k$. Eberhard Hopf showed:

Theorem 1 (Maximal ergodic theorem)**.** *Given $X \in \mathcal{L}^1$ and a measure preserving T, the event $A = \{\omega, \sup_n S_n(\omega) > 0\ \}$ satisfies* $\mathrm{E}[X; A] = \mathrm{E}[1_A X] \geq 0$.

Proof. Define $Z_n = \max_{0 \leq k \leq n} S_k$ and the sets $A_n = \{Z_n > 0\} \subset A_{n+1}$. Then $A = \bigcup_n A_n$. Clearly $Z_n \in \mathcal{L}^1$. For $0 \leq k \leq n$, we have $Z_n \geq S_k$ and so $Z_n(T) \geq S_k(T)$ and hence $Z_n(T) + X \geq S_{k+1}$. By taking the maxima on both sides over $0 \leq k \leq n$, we get $Z_n(T) + X \geq \max_{1 \leq k \leq n+1} S_k$. On $A_n = \{Z_n > 0\}$, we can extend this to $Z_n(T) + X \geq \max_{1 \leq k \leq n+1} S_k \geq \max_{0 \leq k \leq n+1} S_k = Z_{n+1} \geq Z_n$ so that on A_n we have $X \geq Z_n - Z_n(T)$. Integration over the set A_n gives $\mathrm{E}[X; A_n] \geq \mathrm{E}[Z_n; A_n] - \mathrm{E}[Z_n(T); A_n]$ Using (1) this inequality, the fact (2) that $Z_n = 0$ on $\Omega \setminus A_n$, the (3) inequality $Z_n(T) \geq S_n(T) \geq 0$ on A_n and finally that T is measure preserving (4), leads to

$$\begin{aligned} \mathrm{E}[X; A_n] &\geq_{(1)} \mathrm{E}[Z_n; A_n] - \mathrm{E}[Z_n(T); A_n] \\ &=_{(2)} \mathrm{E}[Z_n] - \mathrm{E}[Z_n(T); A_n] \\ &\geq_{(3)} \mathrm{E}[Z_n - Z_n(T)] =_{(4)} 0 \end{aligned}$$

for every n and so to $\mathrm{E}[X; A] \geq 0$. ∎

Theorem 2 (Birkhoff, 1931)**.** *For $X \in \mathcal{L}^1$ and ergodic T, one has $\frac{S_n}{n} \to^{a.e.} \mathrm{E}[X]$.*

Proof. Define $\overline{X} = \limsup_{n\to\infty} \frac{S_n}{n}$, $\underline{X} = \liminf_{n\to\infty} \frac{S_n}{n}$. We get $\overline{X} = \overline{X(T)}$ and $\underline{X} = \underline{X}(T)$ because

$$\frac{n+1}{n} \frac{S_{n+1}}{(n+1)} - \frac{S_n(T)}{n} = \frac{X}{n} \; .$$

(i) $\overline{X} = \underline{X}$. Define for $\beta < \alpha \in \mathbb{R}$ the set $A_{\alpha,\beta} = \{\underline{X} < \beta < \alpha < \overline{X}\ \}$. It is T-invariant because $\overline{X}, \underline{X}$ are T-invariant as mentioned at the beginning of the proof. Because $\{\underline{X} < \overline{X}\ \} = \bigcup_{\beta < \alpha, \alpha, \beta \in \mathbb{A}} A_{\alpha,\beta}$, it is enough to show that $\mathrm{P}[A_{\alpha,\beta}] = 0$ for rational $\beta < \alpha$. The rest of the proof establishes this. In order to use the maximal ergodic theorem, we also define

$$\begin{aligned} B_{\alpha,\beta} &= \{\sup_n (S_n - n\alpha) > 0\ , \sup_n (S_n - n\beta) < 0\} \\ &= \{\sup_n (\frac{S_n}{n} - \alpha) > 0, \sup_n (\frac{S_n}{n} - \beta) < 0\ \} \\ &\supset \{\limsup_n (\frac{S_n}{n} - \alpha) > 0, \limsup_n (\frac{S_n}{n} - \beta) < 0\ \} \\ &= \{\overline{X} - \alpha > 0, \underline{X} - \beta < 0\ \} = A_{\alpha,\beta}\ . \end{aligned}$$

Because $A_{\alpha,\beta} \subset B_{\alpha,\beta}$ and $A_{\alpha,\beta}$ is T-invariant, we get from the maximal ergodic theorem $\mathrm{E}[\overline{X} - \alpha; A_{\alpha,\beta}] \geq 0$ and so

$$\mathrm{E}[\overline{X}; A_{\alpha,\beta}] \geq \alpha \cdot \mathrm{P}[A_{\alpha,\beta}]\ .$$

Because $A_{\alpha,\beta}$ is T-invariant, we we get from (i) restricted to the system T on $A_{\alpha,\beta}$ that $\mathrm{E}[\overline{X}; A_{\alpha,\beta}] = \mathrm{E}[X; A_{\alpha,\beta}]$ and so

(1) $$\mathrm{E}[X; A_{\alpha,\beta}] \geq \alpha \cdot \mathrm{P}[A_{\alpha,\beta}]\ .$$

Replacing X, α, β with $-X, -\beta, -\alpha$ and using $\overline{-X} = -\underline{X}$ shows in exactly the same way that

(2) $$\mathrm{E}[X; A_{\alpha,\beta}] \leq \beta \cdot \mathrm{P}[A_{\alpha,\beta}]\ .$$

The two equations (1),(2) imply that

$$\beta \mathrm{P}[A_{\alpha,\beta}] \geq \alpha \mathrm{P}[A_{\alpha,\beta}]$$

which together with $\beta < \alpha$ only leave us to conclude $\mathrm{P}[A_{\alpha,\beta}] = 0$.

(ii) $\overline{X} \in \mathcal{L}^1$ We have $|S_n/n| \leq |X|$, and by (i) that S_n/n converges point-wise to $\overline{X} = \underline{X}$ and $X \in \mathcal{L}^1$. The Lebesgue's dominated convergence theorem assures $\overline{X} \in \mathcal{L}^1$.

(iii) $\mathrm{E}[X] = \mathrm{E}[\overline{X}]$. Define the T-invariant sets $B_{k,n} = \{\overline{X} \in [\frac{k}{n}, \frac{k+1}{n})\}$ for $k \in \mathbb{Z}, n \geq 1$. Define for $\epsilon > 0$ the random variable $Y = X - \frac{k}{n} + \epsilon$ and call $\tilde{S}_n$ the sums where X is replaced by Y. We know that for n large enough $\sup_n \tilde{S}_n \geq 0$ on $B_{k,n}$. When applying the maximal ergodic theorem applied to the random variable Y on $B_{k,n}$. we get $\mathrm{E}[Y; B_{k,n}] \geq 0$. Because $\epsilon > 0$ was arbitrary, $\mathrm{E}[X; B_{k,n}] \geq \frac{k}{n}\mathrm{P}[B_{k,n}]$. With this inequality,

$$\mathrm{E}[\overline{X}, B_{k,n}] \leq \frac{k+1}{n}\mathrm{P}[B_{k,n}] \leq \frac{1}{n}\mathrm{P}[B_{k,n}] + \frac{k}{n}\mathrm{P}[B_{k,n}] \leq \frac{1}{n}\mathrm{P}[B_{k,n}] + \mathrm{E}[X; B_{k,n}]\ .$$

Summing over k gives

$$\mathrm{E}[\overline{X}] \leq \frac{1}{n} + \mathrm{E}[X]$$

and because n was arbitrary, $\mathrm{E}[\overline{X}] \leq \mathrm{E}[X]$. Doing the same with $-X$ we end with

$$\mathrm{E}[-\overline{X}] = \mathrm{E}[\underline{-X}] \leq \mathrm{E}[\overline{-X}] \leq \mathrm{E}[-X]\ .$$

$\square$