1

- Let $(X, \mathcal{B}, \mu)$ be a measurable space with a probability measure $\mu$. $f$ be a measure-preserving map. Recall

(Kolomogorov-Sinai)

<u>Thm</u>: Let $\mathcal{A}$ be a finite measurable partition of $X$. if $\bigvee_{i=0}^{\infty} f^{-i}\mathcal{A}$ generates the $\sigma$-algebra $\mathcal{B}$. then

$$h_\mu(f) = h_\mu(f, \mathcal{A}).$$

<u>Rk</u>: $\bigvee_{i=0}^{\infty} f^{-i}\mathcal{A} = \bigcup_{k=0}^{\infty} \bigvee_{i=0}^{k} f^{-i}\mathcal{A}$. is Bool algebra

<u>Preparation</u>:

<u>lem</u>: $\forall$ $B \in \mathcal{B}$. and $\varepsilon > 0$. $\exists$ $k$ and. a finite union of subsets ~~as a~~ of $\bigvee_{i=0}^{k} f^{-i}\mathcal{A}$. called $A$. s.t

$$\mu(B \triangle A) < \varepsilon.$$

<u>Pf</u>: By definition of generating $\sigma$-algebra

<u>lem</u>. $\alpha, \beta, \gamma$ three measurable partition

- $H(\alpha \vee \gamma | \beta) = H(\alpha|\beta) + H(\gamma | \alpha \vee \beta)$
- $H(\alpha | \beta \vee \gamma) \leq H(\alpha|\beta)$

Pf: (1) $H(\alpha \vee \gamma | \beta)$

$$= \sum_{i,j,k} -\mu(A_i \cap C_j \cap B_k) \log \frac{\mu(A_i \cap C_j \cap B_k)}{\mu(B_k)}$$

$$= \sum \qquad \left( \log \frac{\mu(A_i \cap C_j \cap B_k)}{\mu(A_i \cap B_k)} + \log \frac{\mu(A_i \cap B_k)}{\mu(B_k)} \right)$$

$$= H(\gamma | \alpha \vee \beta) + \sum_{i,j,k} -\mu(A_i \cap C_j \cap B_k) \log \left( \frac{\mu(A_i \cap B_k)}{\mu(B_k)} \right)$$

$$= H(\gamma | \alpha \vee \beta) + H(\alpha | \beta)$$

(2) $H(\alpha | \beta \vee \gamma)$

$$= \sum_{i,j,k} -\mu(A_i \cap B_k \cap C_j) \log \frac{\mu(A_i \cap B_k \cap C_j)}{\mu(B_k \cap C_j)}$$

$$= \sum_{i} \sum_{k} \mu(B_k) \sum_{j} \left( - \frac{\mu(A_i \cap B_k \cap C_j)}{\mu(B_k)} \log \frac{\mu(A_i \cap B_k \cap C_j)}{\mu(B_k \cap C_j)} \right)$$

Apply Jensen to $t_j = \frac{\mu(B_k \cap C_j)}{\mu(B_k)}$. $x_j = \frac{\mu(A_i \cap B_k \cap C_j)}{\mu(B_k \cap C_j)}$

$$\leq \sum_{k,i} \mu(B_k) \, \varphi\left(\sum_j t_j x_j\right)$$

$$= \sum_{k,i} \mu(B_k) \left( - \frac{\mu(A_i \cap B_k)}{\mu(B_k)} \log \frac{\mu(A_i \cap B_k)}{\mu(B_k)} \right) = H(\alpha | \beta).$$

Lem: Let $\beta_n < \beta_{n+1} < \cdots$ be a sequence of measurable partition s.t. the union generates the $\sigma$-algebra

Then for any ~~measurable~~ finite partition $\alpha$

we have $\lim_{n\to\infty} H_\mu(\alpha \mid \beta_n) = 0$.

Pf: Suppose $\alpha = \{A_1, \cdots, A_p\}$. fix $\varepsilon > 0$

By lem of generating $\sigma$-algebra, $\exists\, n_0$

$\forall A_i$, $\exists$ $B_i$ a finite union of sets in $\beta_{n_0}$

s.t. $\mu(A_i \triangle B_i) < \varepsilon$.

Then for $i \neq j$.

$$\mu(B_i \triangle B_j) \leq 2\varepsilon + \mu(A_i \triangle A_j) = 2\varepsilon$$

(✩)

$$\mu(X - \cup B_j) \leq p\varepsilon + \mu(X - \cup A_j) = p\varepsilon.$$

Hence $\{B_j\}$ is almost a partition.

Then we want modify $B_j$ to $C_j$.

~~s.t.~~ with $C_1 = B_1$ $C_2 = B_2 - C_1$.

$$\cdots\ C_p = (B_p - \bigcup_{j=1}^{p-1} C_j) \cup (X - \cup B_j).$$

By (✩) we know $\mu(A_j \triangle C_j) < C\varepsilon$.

for some $C$ only depending on $p$.

Therefore.

$$\frac{\mu(A_j \cap C_i)}{\mu(C_i)} \text{ is either } \leq \delta \text{ or } \geq 1-\delta.$$

$\delta = \max_{1\leq i\leq p} \left\{ \frac{C\varepsilon}{\mu(A_i)} \right\}$. which is small if $\varepsilon$ small

Hence by lem (ii)

$$H(\alpha \mid \beta_{n_0}) \leq H(\alpha \mid \mathcal{C})$$

$$= \sum_i \mu(C_i) \sum_j - \frac{\mu(A_j \cap C_i)}{\mu(C_i)} \log \frac{\mu(A_j \cap C_i)}{\mu(C_i)}$$

$$\leq \sum_i \mu(C_i)\, p \max\{\varphi(\delta), \varphi(1-\delta)\}$$

$$\leq p \max\{\varphi(\delta), \varphi(1-\delta)\} \xrightarrow{\delta \to 0} 0$$

$\square$

(Rokhlin

lem: Let $\alpha, \beta$ be 2 partitions. then

$$|h_\mu(T,\alpha) - h_\mu(T,\beta)| \leq H(\alpha/\beta).$$

Pf: $H_\mu\left(\bigvee_{i=0}^{n} T^{-i}\alpha\right) - H_\mu\left(\bigvee_{i=0}^{n} T^{-i}\beta\right)$.

$$\leq H_\mu\left(\bigvee_{i=0}^{n} T^{-i}\alpha \vee \bigvee_{i=0}^{n} T^{-i}\beta\right) - H_\mu\left(\bigvee_{i=0}^{n} T^{-i}\beta\right)$$

$$= H_\mu\left(\bigvee_{i=0}^{n} T^{-i}\alpha \,\Big|\, \bigvee_{i=0}^{n} T^{-i}\beta\right)$$

$$\leq \sum_i H_\mu(T^{-i}\alpha \mid \bigvee_{j=0}^{n} T^{-j}\beta)$$

$$\leq \sum_i H_\mu(T^{-i}\alpha \mid T^{-i}\beta) = n H_\mu(\alpha \mid \beta)$$

$\square$

<u>lem</u>: $h_\mu(T,\alpha) = h_\mu(T, \bigvee_{i=0}^{k} T^{-i}\alpha)$

<u>By definition</u>.

<u>Pf of thm</u>. $\forall \varepsilon > 0$ $\exists$ $\mathcal{B}$. s.t.

$$h_\mu(T, \mathcal{B}) > h_\mu(T) - \varepsilon.$$

$\exists\, n_0$. s.t. By generating

$$H_\mu(\mathcal{B} \mid \bigvee_{i=0}^{n_0} T^{-i}\mathcal{A}) < \varepsilon$$

By <u>lem</u>.

$$h_\mu(T,\mathcal{A}) = h_\mu(T, \bigvee_{i=0}^{n_0} T^{-i}\mathcal{A})$$

$$\geq h_\mu(T,\mathcal{B}) - H(\mathcal{B} \mid \bigvee_{i=0}^{n_0} T^{-i}\mathcal{A})$$

$$\geq h_\mu(T) - 2\varepsilon$$

$\square$

---

More property of entropy

Cor: (1) $\frac{1}{n} H_\mu\left( \bigvee_{i=0}^{n-1} T^{-i} A \right)$.

is decreasing

(2) $h_\mu(T, A) = \lim H_\mu\left(A \middle| \bigvee_{i=0}^{n-1} T^{-i} A\right)$.

Pf: (1) $H_\mu\left( \bigvee_{i=0}^{n} T^{-i} A \right) = H_\mu\left(A \middle| \bigvee_{i=1}^{n} T^{-i} A\right) + H_\mu\left( \bigvee_{i=1}^{n} T^{-i} A \right)$

$$= H_\mu\left(A \middle| \bigvee_{i=1}^{n} T^{-i} A\right) + H_\mu\left( \bigvee_{i=0}^{n-1} T^{-i} A \right)$$

$\cdots$

$$= \left( \sum_{k=n}^{1} H_\mu\left(A \middle| \bigvee_{i=1}^{k} T^{-i} A\right) \right) + H_\mu(A)$$

Let $c_k = H_\mu\left(A \middle| \bigvee_{i=1}^{k} T^{-i} A\right)$.

then due to $\bigvee_{i=1}^{k} T^{-i} A$ finer than $\bigvee_{i=1}^{k-1} T^{-i} A$,

we have $c_k \leq c_{k-1}$, a decreasing sequence

Hence the Cesaro average is also decreasing

(2) Due to the series is decreasing $\square$

The average information that you can get when you know the images $f^i(x)$.

# Measure of maximal entropy.

ex: Circle rotation (irrational).

Take ~~as~~ partition $P = [a,b] \cup [a,b]^c$.

then $\bigvee_{i=0}^{\infty} T^{-i} P$ generates the $\sigma$-algebra

Hence $\mu = Leb$

$$h_\mu(T) = h_\mu(T, P) = \lim \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n} T^{-i} P\right)$$

Since $\bigvee_{i=0}^{n} T^{-i} P$ has at most $4n$ elements

( $2n$ different endpoints of interval )

$$H_\mu\left(\bigvee_{i=0}^{n} T^{-i} P\right) \leq \log(4n)$$

$$\Rightarrow \quad h_\mu(T) = 0$$

ex: Times ~~p~~ $p$ map on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$.

Take a partition $P = \bigcup_{j=0}^{p-1} [\frac{j}{p}, \frac{j+1}{p})$.

Then $T_p^{-(n-1)} P = \bigcup [\frac{j}{p^n}, \frac{j+1}{p^n})$,

~~id~~ $\bigvee_{i=0}^{\infty} T_p^{-i} P$ generates the Borel- algebra.

Hence $h_\mu(T) = h_\mu(T, P) = \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n} T^{-i} P\right)$

For $\mu =$ Leb. we have

$$H_{Leb}\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) = \sum_j - p^{-n} \log p^{-n} = n \log p$$

Hence $h_{Leb}(T) = \log p$.

ex. Bernoulli shift on $p$-letters $\Lambda_p = \{1, \cdots, p\}$

Due to measurable isomorphism, we also know

$(\Lambda_p^{\otimes \mathbb{N}}, \text{shift}, \overset{\mathbb{N}}{\otimes} \nu)$ has entropy $\log p$.

If $\nu = \sum q_i \delta_i$. then

$$h_\mu(T) = \sum - q_i \log q_i$$

Orstein: Entropy is a complete invariant for Bernoulli shift.

Measure of maximal entropy:

For times $p$ map., for any invariant measure $\mu$. we have

$$h_\mu(T) = \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\right) \le \lim \frac{1}{n} \log p^n$$
$$= \log p$$

Hence Lebesgue measure is of maximal entropy. 9

<u>Thm</u>: Lebesgue measure is the unique measure with maximal entropy.

<u>Pf</u>. Let $\mu$ be another measure with maximal entropy than Leb

First notice $\mu$ is singular to Leb, otherwise by Radon-Nykodym $d\mu = f\, d\,\mathrm{Leb}$
+ $\mu$ is $T$-inv. and Leb is ergodic $\Rightarrow f = cst$.

Since $\mu$ is singular, $\exists$ a Borel set $B$. s.t.

$$\mu(B) > 0 \quad \text{but} \quad \nu(B) = 0$$

Take $\varepsilon < \frac{1}{2}\mu(B)$. Due to $\bigvee_0^\infty T^{-i}\mathcal{P}$ generating $\sigma$-algebra, $\exists$ ~~$n_0$~~ $n$ and a set $B_n$ as a finite union of atoms in $\bigvee_0^n T^{-i}\mathcal{P}$.

s.t. $\mu(B_n) > \mu(B) - \varepsilon$, $\quad \nu(B_n) < \varepsilon$

Then $\forall$ $m \geq n$. we have

(1) $$\frac{1}{m} H_\mu\left(\bigvee_0^m T^{-i}\mathcal{P}\right) = \frac{1}{m} H_\mu\left(\bigvee_0^m T^{-i}\mathcal{P} \,\middle|\, B_n \cup B_n^c\right) + \frac{1}{m} H_\mu(B_n \cup B_n^c)$$

Since $\bigvee_0^m T^{-i}\mathcal{P}$ finer than $\{B_n, B_n^c\}$

Due to corollary, we have (Cor (1)).

$\frac{1}{m} H_\mu(\bigvee_0^m T^{-i}\beta)$ decreasing, so

(2) $\frac{1}{m} H_\mu(\bigvee_0^m T^{-i}\beta) \geq \log p$

(3) $H_\mu(\bigvee_0^m T^{-i}\beta / B_n \cup B_n^c)$

$= \mu(B_n) H_{\mu/B_n}(\bigvee_0^m T^{-i}\beta) + \mu(B_n^c) H_{\mu/B_n^c}(\bigvee_0^m T^{-i}\beta)$

$\leq \mu(B_n) \log (\nu(B_n)) p^m + \mu(B_n^c) \log (1-\nu(B_n)) p^m.$

$= \mu(B_n) \log \nu(B_n) + \mu(B_n^c) \log \nu(B_n^c) + m \log p$

By (1), (2), (3)

$0 \leq \cancel{\frac{1}{m}} H_\mu(\bigvee_0^m T^{-i}\beta) - m\log p$

$\leq \mu(B_n) \log \nu(B_n) + \mu(B_n^c) \log \nu(B_n^c) + H_\mu(\{B_n, B_n^c\})$

$= \mu(B_n) \log \frac{\nu(B_n)}{\mu(B_n)} + \underline{\mu(B_n^c) \log \frac{\nu(B_n^c)}{\mu(B_n^c)}}$

$\longrightarrow -\infty$ bounded.

Contradiction