More Entropy, Atomicity, and Uniformity

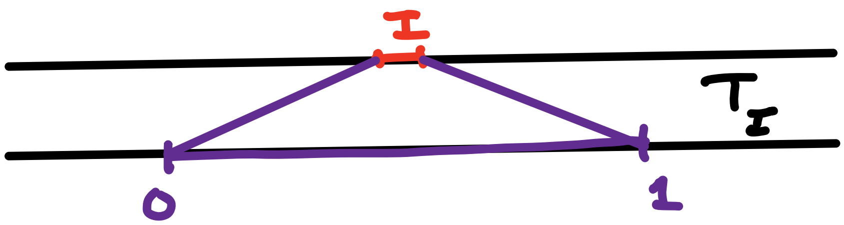
Let $\mathcal{E}$ be a countable partition of $\mathbb{R}$. Let $E \in \mathcal{E}$.

For a Borel measure, μ, on $\mathbb{R}$ denote

$$\mu_E := \frac{1}{\mu(E)} \mu|_E$$ (If $\mu(E)>0$)

For an interval, I, let $T_I : \mathbb{R} \to \mathbb{R}$ be the homothety map with

$$T_I(I) = [0,1)$$



define

$$\mu^I := \frac{1}{\mu(I)} (T_I)_\# (\mu|_I)$$

Recall that

$$H(\mu, \mathcal{E}) = -\sum_{E \in \mathcal{E}} \mu(E) \log(\mu(E))$$

for another partition, $\mathcal{F}$, define the <u>conditional entropy</u> as

$$H(\mu, \mathcal{E} | \mathcal{F}) = \sum_{F \in \mathcal{F}} \mu(F)\, H(\mu_F, \mathcal{E})$$

$$= \sum_{F \in \mathcal{F}} \mu(F) \Big(\sum_{E \in \mathcal{E}} \mu_F(E) \log(\mu_F(E)) \Big)$$

Finally, recall that for a probability vector $p = (p_2, \ldots, p_m)$

$$H(p) = \sum_{i=1}^{m} p_i \log(p_i)$$

<u>Notes</u>

- $\mathcal{E} \vee \mathcal{F} = \{ E \cap F \mid F \in \mathcal{F}, E \in \mathcal{E} \}$

  $$H(\mu, \mathcal{E} \vee \mathcal{F}) = H(\mu, \mathcal{F}) + H(\mu, \mathcal{E} | \mathcal{F})$$

[Very important identity. Will be used in different forms: If $\mathcal{E}$ is a refinement of $\mathcal{F}$, then

$$H(\mu, \mathcal{E} | \mathcal{F}) = H(\mu, \mathcal{E}) - H(\mu, \mathcal{F}).$$].

- For any prob. vector $p=(p_1,\dots,p_m)$

$$\sum_{i=1}^{m} p_i H(\mu_i, \mathcal{E}) \le H\left(\sum_{i=1}^{m} p_i \mu_i, \mathcal{E}\right)$$
$$\le \sum_{i=1}^{m} p_i H(\mu_i, \mathcal{E}) + H(p).$$

Lemma:

Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ = prob measures on $\mathbb{R}$, $\mathcal{E}, \mathcal{F}$ countable partitions on $\mathbb{R}$ and let $m, n \in \mathbb{Z}_+$

a.) If $m<n$, $H(\mu, \mathcal{D}_m) \le H(\mu, \mathcal{D}_n)$

b.) $H(\mu, \mathcal{D}_n) \le n$, $H(\mu, \mathcal{D}_n | \mathcal{D}_m) \le n-m$

c.) Let $K \subset \mathbb{R}$ be compact, $\mu \in \mathcal{P}(K)$. there exist a nbhd of measures $U \subset \mathcal{P}(K)$ s.t.

$$\left| H(\mu, \mathcal{D}_m) - H(\nu, \mathcal{D}_m) \right| \le 1 \quad \forall \nu \in U.$$

d.) If each $E \in \mathcal{E}$ intersects at most k many $F \in \mathcal{F}$ and if each $F \in \mathcal{F}$ intersects at most k many $E \in \mathcal{E}$, then

$$\left| H(\mu, \mathcal{E}) - H(\mu, \mathcal{F}) \right| \lesssim \log(k).$$

e.) If $f, g : \mathbb{R} \to \mathbb{R}$ and $\sup_x |f(x) - g(x)| \le 2^{-m}$ then

$$|H(f_* \mu, \mathcal{D}_m) - H(g_* \mu, \mathcal{D}_m)| \lesssim 1.$$

f.) If $\nu(\cdot) = \mu(\cdot + x_0)$, then

$$|H(\mu, \mathcal{D}_m) - H(\nu, \mathcal{D}_m)| \lesssim 1$$

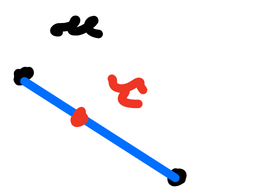
g.) If $C^{-1} \le \frac{m'}{m} \le C$, then

$$|H(\mu, \mathcal{D}_m) - H(\mu, \mathcal{D}_{m'})| \lesssim_C 1.$$

Let $\|\mu - \nu\| := \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)|$

For every $\varepsilon > 0$, $\exists \delta > 0$ s.t. if $\|\mu - \nu\| < \delta$, then $\exists\, \tau, \mu', \nu' \in \mathcal{P}(\mathbb{R})$ such that

$$\mu = (1-\varepsilon)\tau + \varepsilon \mu'$$
$$\nu = (1-\varepsilon)\tau + \varepsilon \nu'$$



$\tau = \frac{1}{2}(\mu + \nu)$

$$\mu = (1-\varepsilon)\tau + \varepsilon\left(\frac{1+\varepsilon}{2\varepsilon}\mu - \frac{1-\varepsilon}{2\varepsilon}\nu\right).$$

This leads to the following Lemma

Lem:

For every $\varepsilon > 0$, $\exists \delta > 0$ s.t.

if $\mu, \nu \in \mathcal{P}(\mathbb{R})$ and $\|\mu - \nu\| < \delta$

then if $\mathcal{E}$ is a finite partition

with k elements then

$$\left| H(\mu, \mathcal{E}) - H(\nu, \mathcal{E}) \right| < \varepsilon \log(k) + \varepsilon \log(\varepsilon).$$

If we let

$$H_m(\mu) := \frac{1}{m} H(\mu, \mathcal{D}_m)$$

and $\mu, \nu \in \mathcal{P}([0,1])$

then

$$\left| H_m(\mu) - H_m(\nu) \right| < \varepsilon + \frac{\varepsilon \log(\varepsilon)}{m}$$

Probabalistic Entropy

For $n \in \mathbb{Z}_+$, $x \in \mathbb{R}$

Let $q(x) \in \mathcal{D}_n$ be the unique dyadic interval containing x.

If $\mu(q(x)) > 0$, then

$$\mu_{x,n} := \mu_{q(x)} = \frac{1}{\mu(q(x))} \cdot \mu|_{q(x)}.$$

and

$$\mu^{x,n} := \mu^{q(x)} = \frac{1}{\mu(q(x))} (T_{q(x)})_* \left(\mu|_{q(x)}\right)$$

We can think of $\mu^{x,i}$ and $\mu_{x,i}$ as random variables of measures

$$\mu^{\cdot,\cdot} : \mathbb{R} \times \{0, \dots, n\} \to \mathcal{P}(\mathbb{R})$$

with

$$\mathbb{P}_{i=n}(\mu^{x,i} \in A) = \int \chi_A(\mu^{x,n}) \, d\mu(x)$$

$$\left(\begin{array}{ll} & \mathbb{P}_{i=n}(\mu^{x,i} = \mu^q) = \int \chi_{\{\mu^q\}}(\mu^{x,n}) \, d\mu(x) \\ \text{If } q \in \mathcal{D}_n & = \int \chi_q(x) \, d\mu(x) = \mu(q) \end{array} \right)$$

$$\mathbb{P}_{0\le i\le n}(\mu^{x,i}\in A)$$

$$= \frac{1}{n+1}\sum_{i=0}^{n}\int \chi_A(\mu^{x,i})\, d\mu(x)$$

Example:

Let $q\in\Delta_k$, let $\mu=\frac{1}{\mathcal{L}^1(q)}\left(\mathcal{L}^1\big|_q\right)$
let $\lambda=\mathcal{L}^1\big|_{[0,1]}$

$$\mathbb{P}_{0\le i\le n}(\mu^{x,i}\in\{\lambda\})$$

$$=\frac{1}{n+1}\sum_{i=0}^{n}\int \chi_{\{\lambda\}}(\mu^{x,i})\, d\mu(x)$$

For $i<k$, $\mu^{x,i}\neq\lambda$ for all x.

For $k\le i\le n$, $\mu^{x,i}=\lambda$ for $x\in q$.

Thus

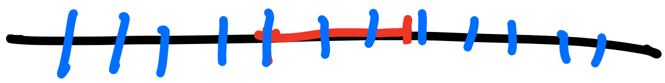
$$\mathbb{P}_{0\le i\le n}(\mu^{x,i}\in\{\lambda\})=\frac{1}{n+1}\sum_{i=k}^{n}\mu(q)$$

$$=\frac{n-k}{n+1}$$

(This is similar to computing how many steps are necessary to describe a set q with a Turing machine)

Note that

$$H(\mu^{x,n}, \mathcal{D}_m) = H(\mu_{x,n}, \mathcal{D}_{n+m})$$

and

$$\mathbb{E}_{i=n}\left(H_m(\mu^{x,i})\right)$$

$$= \int \frac{1}{m} H(\mu^{x,n}, \mathcal{D}_m)\, d\mu(x)$$

$$= \int \frac{1}{m} H(\mu_{x,n}, \mathcal{D}_{n+m})\, d\mu(x)$$

$$= \frac{1}{m} \sum_{Q \in \mathcal{D}_n} \mu(Q)\, H(\mu_Q, \mathcal{D}_{m+n})$$

$$= \frac{1}{m} H(\mu, \mathcal{D}_{n+m} \mid \mathcal{D}_n)$$

Lemma:

For $r \geq 1$, $\mu \in \mathcal{P}([-r,r])$ and $m < n$,

$$H_n(\mu) = \mathbb{E}_{0 \leq i \leq n}\left(H_m(\mu^{x,i})\right) + O\left(\frac{m}{n} + \frac{\log r}{n}\right)$$

"Global entropy can be determined by local entropy"

pf:

Since $\mathcal{D}_m \vee \mathcal{D}_n = \mathcal{D}_n$ for $m < n$,

thus

$$H(\mu, \mathcal{D}_n | \mathcal{D}_m) = H(\mu, \mathcal{D}_m \vee \mathcal{D}_n) - H(\mu, \mathcal{D}_m)$$

$$= H(\mu, \mathcal{D}_n) - H(\mu, \mathcal{D}_m).$$

Observe that

$$\mathbb{E}_{0 \leq i \leq n}[H_m(\mu^{x,i})]$$

$$= \frac{1}{n+1}\sum_{i=0}^{n} \mathbb{E}_{j=i}[H_m(\mu^{x,j})]$$

$$= \frac{1}{n+1}\sum_{i=0}^{n} \frac{1}{m} H(\mu, \mathcal{D}_{i+m} | \mathcal{D}_i)$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{m} H(x, D_{i+m} \mid D_i) + O\left(\frac{n-m}{n\,m} \cdot \log(r)\right)$$

So it suffices to show that

$$H_n(x) = \frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{m} H(x, D_{i+m} \mid D_i) + O\left(\frac{m}{n} + \frac{\log(r)}{m}\right).$$

For $i=1$, we simply have

$$H_n(x) = \frac{1}{n} H(x, D_n)$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} H(x, D_{i+1} \mid D_i) + \frac{1}{n} H(x, D_0)$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} H(x, D_{i+1} \mid D_i) + O\left(\frac{\log r}{n}\right)$$

and we are done

Otherwise,

$$H(x, D_n) = \sum_{i=0}^{m-1} \frac{1}{m} H(x, D_{n+i}) - \frac{1}{m} \sum_{i=0}^{m-1} H(x, D_{n+i} \mid D_n)$$

$$= \sum_{i=0}^{m-1} \frac{1}{m} H(x, D_{n+i}) + O(m)$$

For each i

$$H(\mu, \mathcal{D}_{n+i}) = \sum_{j=1}^{\lfloor \frac{n}{m} \rfloor + 1} H(\mu, \mathcal{D}_{n+i-(j-1)m} | \mathcal{D}_{n+i-jm}) + H(\mu, \mathcal{D}_e) \quad O(m)$$

Combining the same for all i, we get

$$H(\mu, \mathcal{D}_n) = \sum_{i=0}^{n} \frac{1}{m} H(\mu, \mathcal{D}_{i+m} | \mathcal{D}_i)$$

$\square$

Similarly, we have the following lemma.

<u>Lemma:</u> Let $\mu, \nu \in \mathcal{P}([-r, r])$ and $m < n$. Then

$$H_n(\mu * \nu) \geq \mathbb{E}_{0 \leq i \leq n} \left[H(\mu^{x,i} * \nu^{x,i}) \right] + O\left(\frac{1}{m} + \frac{m}{n} + \frac{\log(r)}{n}\right)$$