A local continuity bound for the entropy of finite distributions

Eric P. Hanson
supervised by Nilanjana Datta

May 24, 2017

Department of Applied Math and Theoretical Physics
Cambridge University

Entropy

Consider the finite set $\Omega = \{1, 2, \ldots, d\}$.

Let p be a probability distribution on Ω: $p(i) \geq 0$, $\sum_{i=1}^{d} p(i) = 1$.
We will call the set of all probability distributions on Ω by Δ.

Then the *entropy* of p is

$$H(p) := -\sum_{i=1}^{d} p(i) \log p(i).$$

- Measures information: how much to learn from observing an outcome
- Min: $(1, 0, \ldots, 0)$. Nothing to learn, zero entropy.
- Max: $(1/d, \ldots, 1/d)$. Most uncertainty.

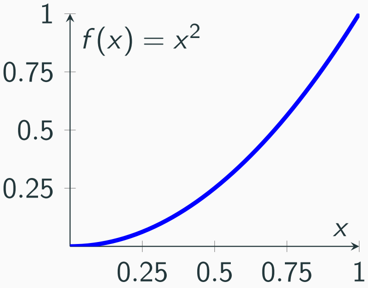
Why entropy?

- Fundamental quantity in information theory
  - Shannon source coding. Given i.i.d. data distributed with probability distribution p, the asymptotic optimal lossless data-compression rate is $H(p)$.

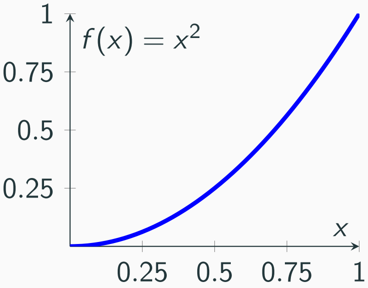
Why entropy?

- Fundamental quantity in information theory
  - Shannon source coding. Given i.i.d. data distributed with probability distribution p, the asymptotic optimal lossless data-compression rate is $H(p)$.
- Related to other quantities (e.g. conditional entropy when you don't condition on anything)
- The von Neumman entropy of a d-dimensional quantum state ρ is the (Shannon) entropy of its eigenvalues, which form a probability distribution on Ω.

Continuity bounds



Continuity bounds



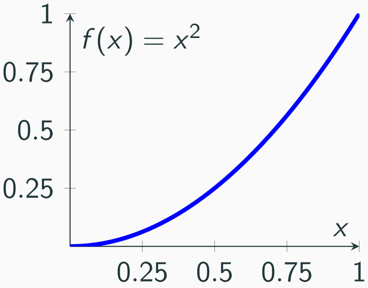
Continuity bound: don't know the values x, y but know

- $x, y \in [0, 1]$
- $|x - y| \leq \varepsilon$.

Continuity bound:

$$|f(x) - f(y)| \ \leq \ 2\varepsilon$$

Continuity bounds



Continuity bound: don't know the values x, y but know

- $x, y \in [0,1]$
- $|x - y| \leq \varepsilon$.

Continuity bound:

$$|f(x) - f(y)| \;\leq\; 2\varepsilon$$

Local continuity bound: do know the value x. LCB:

$$|f(x) - f(y)| \leq (x+1)\varepsilon.$$

Tighter bound for any $x < 1$.

Audenaert-Fannes bound

There already exists a continuity bound for entropy!

Lemma (Audenaert-Fannes)

Let $\varepsilon \in [0,1]$*, and* $p, q \in \Delta$ *such that* $\frac{1}{2}\|p-q\|_1 \leq \varepsilon$*. Then*

$$|H(p) - H(q)| \leq \begin{cases} \varepsilon \log(d-1) + h(\varepsilon) & \textit{if } \varepsilon < 1 - \frac{1}{d} \\ \log d & \textit{if } \varepsilon \geq 1 - \frac{1}{d}. \end{cases} \tag{AF}$$

LCB from maximal and minimal distributions

$$\boxed{B_\varepsilon(q) = \{p \in \Delta : \frac{1}{2}\|p - q\|_1 \leq \varepsilon\}}$$

Fix $q \in \Delta$ and $\varepsilon \in [0,1]$. We always have

$$H(p) - H(q) \leq \max_{p \in B_\varepsilon(q)} H(p) - H(q)$$

LCB from maximal and minimal distributions

$$\boxed{B_\varepsilon(q) = \{p \in \Delta : \frac{1}{2}\|p - q\|_1 \le \varepsilon\}}$$

Fix $q \in \Delta$ and $\varepsilon \in [0, 1]$. We always have

$$H(p) - H(q) \le \max_{p \in B_\varepsilon(q)} H(p) - H(q)$$

and similarly for $H(q) - H(p)$, so

$$|H(p) - H(q)| \le \max\Big\{ \max_{p \in B_\varepsilon(q)} H(p) - H(q), H(q) - \min_{p \in B_\varepsilon(q)} H(p)\Big\}$$

which is a function of q and ε only.

LCB from maximal and minimal distributions

$$\boxed{B_\varepsilon(q) = \{p \in \Delta : \frac{1}{2}\|p - q\|_1 \leq \varepsilon\}}$$

Fix $q \in \Delta$ and $\varepsilon \in [0,1]$. We always have

$$H(p) - H(q) \leq \max_{p \in B_\varepsilon(q)} H(p) - H(q)$$

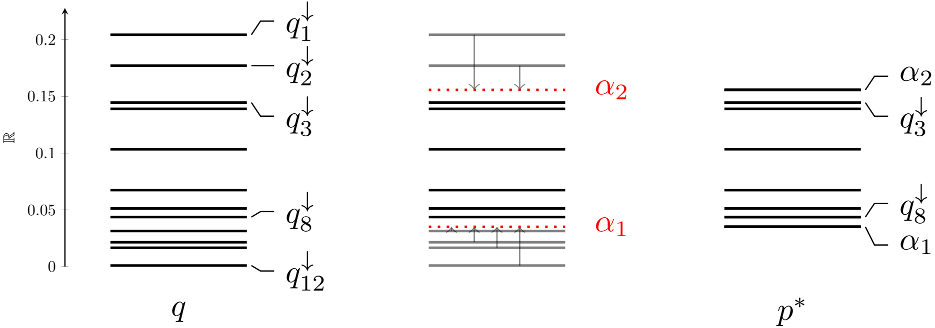
and similarly for $H(q) - H(p)$, so

$$|H(p) - H(q)| \leq \max\Big\{ \max_{p \in B_\varepsilon(q)} H(p) - H(q), H(q) - \min_{p \in B_\varepsilon(q)} H(p) \Big\}$$

which is a function of q and ε only.

Idea: understand $p^* = \mathrm{argmax}_{p \in B_\varepsilon(q)} H(p)$ and $\mathrm{argmin}_{p \in B_\varepsilon(q)}(p)$.

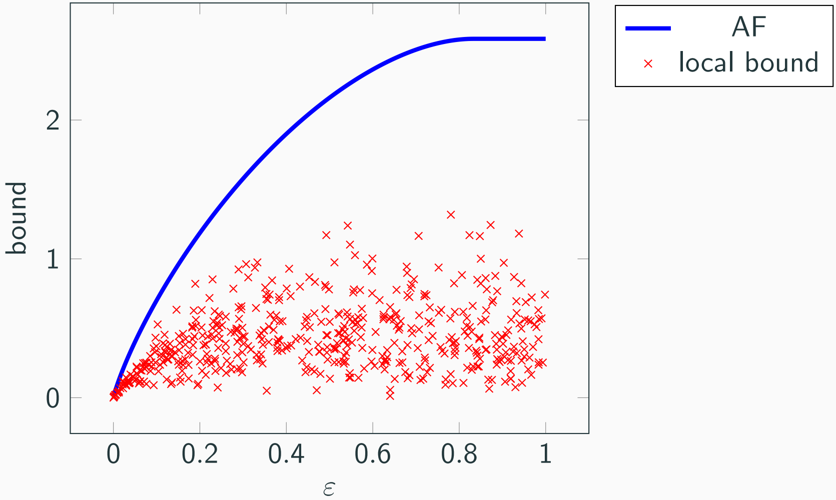
Convex optimization $\rightsquigarrow$ optimality conditions



$\mathbb{R}$
0.2
0.15
0.1
0.05
0
$q_1^{\downarrow}$
$q_2^{\downarrow}$
$q_3^{\downarrow}$
$q_8^{\downarrow}$
$q_{12}^{\downarrow}$
q
α_2
α_1
α_2
$q_3^{\downarrow}$
$q_8^{\downarrow}$
α_1
p^*

Comparison of local continuity bound and the AF bound

In red: 500 choices of $q \in \Delta$ and $\varepsilon \in [0, 1]$.



Thank you!