15-859 ITCS | 3-3 problem sets

lecture 1 | Scribe notes

Let X be a discrete r.v.

$$X \leftarrow \begin{array}{cc} a & 2/3 \\ b & 1/6 \\ c & 1/6 \end{array}$$

$H(X)$: entropy of r.v. X

measures information conveyed by the value of X.

$S : [0,1] \to \mathbb{R}^+$ "surprise measure"

$S(1) = 0$

$S(p) > S(q)$ if $p < q$.

$S(x)$ is a continuous function

X_1, X_2 two independent instantiations of X.

$$S\left(\frac{2}{3} \cdot \frac{1}{3}\right) = S\left(\frac{2}{3}\right) + S\left(\frac{1}{3}\right)$$

$\Rightarrow$ Only S satisfying above is

$$S(p) = c \log_2 \left(\frac{1}{p}\right), \quad c > 0.$$

also, $S(\overset{1/2}{\text{coin}}) = 1 \Rightarrow c = 1$.

<u>Def.</u> (Entropy)

Entropy of X, denoted $H(X)$,

= average surprise for an outcome

$$= \mathop{\mathbb{E}}_{x \leftarrow X}\left(\log_2 \frac{1}{p(X=x)} \right)$$

$$= \sum_{\substack{x \in \\ \text{Supp}(X)}} p(x) \log_2 \frac{1}{p(x)}.$$

$$X \quad \begin{array}{cc} a_1 & p_1 \\ \vdots & \vdots \\ a_n & p_n \end{array}$$

$$\Rightarrow H(X) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i}, \quad \left(0 \log \frac{1}{0} := 0\right)$$

Ex: (above) $H(X) = \frac{2}{3}\log\frac{3}{2} + \frac{1}{6}\log 6 + \frac{1}{6}\log 6.$

Fact $H(X) \geq 0$, $H(X)=0$ iff X is deterministic

Ex: Suppose X is uniform on $\{a_1 \ldots a_n\}$

$\Rightarrow H(X) = \log_2 n.$

Fact: If $|\mathrm{Supp}(X)| = n$

$$\Rightarrow H(X) \leq \log_2 n$$

$X \leftarrow \mathrm{Ber}(\frac{1}{2})$: $X = \begin{cases} 0 & \text{w.p. } \frac{1}{2} \\ 1 & \text{''} \end{cases}$

$H(X) = 1$ bit

$X \leftarrow \mathrm{Ber}(1)$: $H(0) = 0$.

$X \leftarrow \mathrm{Ber}(p)$: $\begin{cases} 0 & \text{w.p. } 1-p \\ 1 & \text{w.p. } p \end{cases}$

$X_1 X_2 \ldots X_n$ iid samples

Suppose $p = .0001$. How ~~to~~ compress?

$*\begin{cases} \text{Send } k \text{ (\# of 1's)} \\ \text{Then } \left\lceil \log_2 \binom{n}{k} \right\rceil \text{ bits to reveal which} \end{cases}$

$$\mathbb{E}[\#\text{ bits}] = \lceil \log_2 n \rceil + \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot \left\lceil \log_2 \binom{n}{k} \right\rceil$$

$\Rightarrow$ amortized # bits per $X_i \equiv \, m/n$, $n \to \infty$.

Stirling approx: If k is far from pn, the expression $\to 0$.

the non-vanishing term is

$$h(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}.$$

$*$

$$X \leftarrow \{a_1 - a_n\}. \qquad H(X) = ?$$

$$Z_1 := \begin{cases} 1 & \text{if } X = a_1 \\ 0 & \text{else}. \end{cases}$$

$$Z_2 = \begin{cases} a_2 & \frac{p_2}{1-p_1} \\ \vdots & \vdots \\ a_n & \frac{p_n}{1-p_1} \end{cases}$$

intuitively, $H(X) = H(Z_1) + \Pr(Z_1=0).\, H(Z_2)$

$$= p_1 \log \frac{1}{p_1} + (1-p_1) \log \frac{1}{p_1}$$

$$+ (1-p_1) \sum_{i=2}^{n} \frac{p_i}{1-p_1} \log \frac{1-p_1}{p_i}$$

$$= \sum_{i=1}^{n} p_i \log \frac{1}{p_i} \qquad \square$$

Lema: $|Supp(X)| = n \Rightarrow H(X) \leq \log_2$

* **Jensen's Inequality:**

$$\mathbb{E}[f(X)] \leq f(\mathbb{E}(X))$$

if f is concave.

e.g., $$\begin{cases} \mathbb{E}[X^2] \geq \mathbb{E}(X)^2 \\ \mathbb{E}[\sqrt{X}] \leq \sqrt{\mathbb{E}[X]} \end{cases}$$

Proof of Lemma

$$H(X) = \mathbb{E}_{x \leftarrow X}\left[\log \frac{1}{p(x)}\right] \quad , f(x) := \log$$

$$\leq \log\left(\mathbb{E}_{x \leftarrow X} \frac{1}{p(x)}\right) = \log n.$$

Data Compression (~~S~~Source Coding)

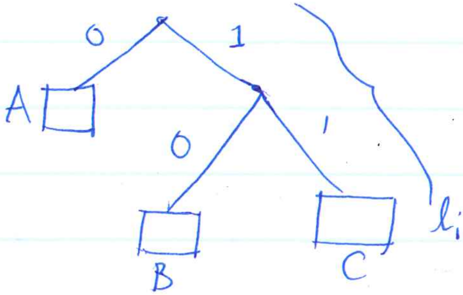
or Lossless Data Compression

$X \leftarrow \{a, b, \dots, z\}$

→ Prefix-free Code or Instantaneous Code

$$\begin{cases} A \to 0 \\ B \to 1 \\ C \to \text{:(} \end{cases} \qquad \begin{cases} A \to 0 \\ B \to 10 \\ C \to 11 \quad \text{:)} \end{cases}$$

$$X \leftarrow \begin{array}{cc} & \text{w.p.} \\ a_1 & p_1 \\ | & | \\ a_n & p_n \end{array}$$

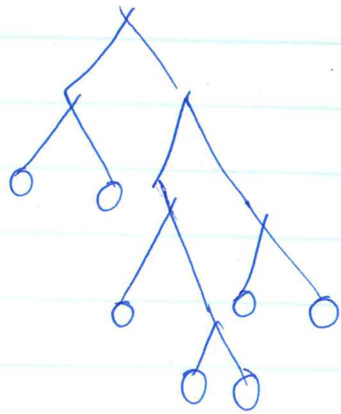


Huffman Coding → Greedy alg.

avg. depth of a symbol $= \sum p_i \, l_i$

Kraft's Inequality: a prefix-free code for n symbols exists of lengths $l_1 \longrightarrow l_n$

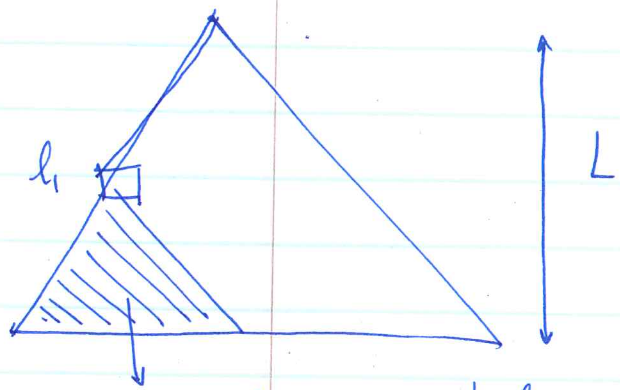
Proof: (~~only~~ if)



random walk from root until fall off or hit a leaf.

Pr (hit a leaf) ≤ 1. $\Rightarrow$ ✓.

($\Rightarrow$) $l_1 \leq l_2 \leq \cdots \leq l_n = L$.



Size $\left(2^{L-l_1}\right) \Rightarrow \dfrac{2^{L-l_1}}{2^L} = 2^{-l_1}$ fraction

& Continue assigning + killing $\longrightarrow 2^{-l_2}$

& never get stuck ✓

(next time: Huffman)