Subject : Intro. to Info. Theory

No. : 1
Date : 9 / 24 / 2019

* probabilistic pigeon hole principle

$0 < \delta < 1$ if $H_\delta(\bar{X}) > H_0(\bar{Y})$

there is no encoder ν & decoder μ

s.t. $\Pr(\mu(\nu(x)) = x) \geq 1-\delta$

ν : deterministic $\bar{X} \to \bar{Y}$

μ : deterministic $\nu(\bar{X}) \to \bar{X}$

> cannot name (specify) $\bar{X}$ w/ fewer than $H_\delta(\bar{X})$ bits (deterministically) to achieve error $\leq \delta$

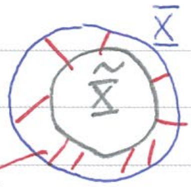
proof : (by contradiction)

assume ν & μ exist while $H_\delta(\bar{X}) > H_0(\bar{Y})$

$\bar{X} \to \nu(\bar{X}) \to \mu(\nu(\bar{X}))$ sufficiently correct

usually ($\geq 1-\delta$) $x \to \nu(x) \to \mu(\nu(x)) = x$

call such $x \in \tilde{\bar{X}}$ always correctly specified



$\leq \delta$ by assumption

$\Downarrow$

$H_\delta(\bar{X}) \leq \log_2(\tilde{\bar{X}})$

$\Downarrow$

$|\tilde{\bar{X}}| \leq |\bar{Y}|$ by deterministic pigeon hole

$\Downarrow$

$H_\delta(\bar{X}) \leq H_0(\bar{Y})$ $\rightarrow\leftarrow$

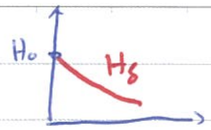
Subject : Intro. to Info. Theory

No. : 2

Date : 9 / 24 / 2019

* $\mathbb{X}$ & p $\xrightarrow{\text{deterministic}}$ $H_0(\mathbb{X})$

$\xrightarrow{\text{tolerance } \delta}$ $H_\delta(\mathbb{X})$



$\longrightarrow$ $H(\mathbb{X})$ "a summary of H_δ w/o depending on δ"

Shannon entropy

* $\mathbb{X} = \underbrace{a_1, a_2, \ldots, a_K}$

a_1 $p(a_1) \to p_1$ as if needing $\log \frac{1}{p(a_1)}$ bits of informati

a_2 $p(a_2) \to p_2$ as if needing $\log \frac{1}{p(a_2)}$ bits of informa

$\vdots$ $\vdots$ $\vdots$ $\vdots$ $\vdots$

a_K $p(a_K) \to p_K$ $\vdots$ $\log \frac{1}{p(a_K)}$ $\vdots$

$$H(\mathbb{X}) = \underbrace{\sum p(a_k)}_{\text{self-avg.}} \underbrace{\log \frac{1}{p(a_k)}}_{\text{ind. info.}}$$

will take $0 \log \frac{1}{0} = 0$

* similar to H_0, certainty = 0 information

$H_0(\{a\}) = 0$

$H(\mathbb{X}) = 0$ iff some $p_i = 1$

* similar to H_δ, $H(\mathbb{X}) \leq H_0(\mathbb{X})$

$$\sum p_k \log \frac{1}{p_k} - \log K = \frac{1}{\ln 2} \sum p_k \boxed{\ln \frac{1}{p_k \cdot K}}$$

$$\leq \frac{1}{\ln 2} \sum p_k \left(\frac{1}{p_k \cdot K} - 1 \right)$$

$$= 0$$

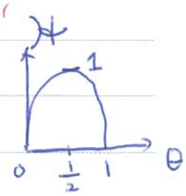
"=" if $\frac{1}{p_k \cdot K} = 1$

* binary ensemble $\mathbb{X} = \{0, 1\}$, $p_0 = \theta$, $p_1 = (1-\theta)$

$$H(\mathbb{X}) = \theta \log \frac{1}{\theta} + (1-\theta) \log \frac{1}{1-\theta} = \mathcal{H}(\theta)$$

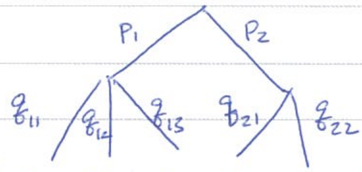
$\mathcal{H}(0) = 0 + 0 = 0$

$\mathcal{H}(1) = 0$



$\max_\theta \mathcal{H}(\theta) =$ "find $\mathcal{H}'(\theta) = 0$" $= \mathcal{H}(0.5) = 1 = H_0(\mathbb{X})$

* additivity of H



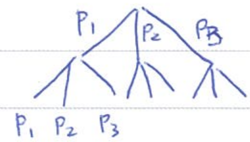
$$\underline{X} = \{ a_{11}, a_{12}, a_{13}, a_{21}, a_{22} \}$$

$$H(\underline{X}) = \sum_i \sum_j p_i q_{ij} \log (p_i q_{ij})$$

$$= \sum_i p_i \sum_j q_{ij} \log q_{ij} + \sum_i p_i \underbrace{\sum q_{ij}}_{=1} \log p_i$$

$$= \sum_i p_i \underbrace{H(\underline{X}_i)}_{\text{sub-level choices}} + \underbrace{H(\underline{X}_0)}_{\text{top-level choice}}$$

* multiple independent experiments



$$H(\underbrace{\underline{X} \cdot \underline{X}})$$

space : $\{ (a_1, a_1), (a_1, a_2), \cdots, (a_K, a_K) \}$

prob : $p_1 \cdot p_1$ $p_1 \cdot p_2$ $p_K \cdot p_K$

$$= H(\underline{X}) + H(\underline{X})$$

$$\Rightarrow H(\underline{X}^N) = N H(\underline{X})$$