

THEORIE DE L'INFORMATION ONE-SHOT

Def: X, Y : VA, distribution $p_{X,Y}$, $\alpha \in (0,1) \cup (1,\infty)$

$$2^{(1-\alpha)H_\alpha(X)} = E_X \left[2^{(1-\alpha)S(X)} \right]$$

$\alpha < 1$: On donne plus de poids aux grandes surprises

$\alpha > 1$: On donne plus de poids aux petites surprises.

$$2^{\frac{1-\alpha}{\alpha} H_\alpha(X|Y)} = E_Y \left[2^{\frac{1-\alpha}{\alpha} H_\alpha(X|Y=Y)} \right]$$

$p, q \in \mathcal{P}(X)$:

$$2^{(\alpha-1)D_\alpha(p||q)}$$

$$= E_{X \sim p} \left[2^{(1-\alpha)(S(X)_p - S(X)_q)} \right]$$

où $H_\alpha(X|Y=Y)$ une VA qui vaut $H_\alpha(X|Y=y)$ quand $Y=y$.

$$2^{\frac{\alpha-1}{\alpha} I_\alpha(X;Y)} = E_Y \left[2^{\frac{\alpha-1}{\alpha} D_\alpha(p_{X|Y=Y} || p_X)} \right]$$

$\hookrightarrow$ PAS symétrique entre X et Y .

Propriétés des mesures de Rényi

- Entropie relative de Rényi est croissante en α :

si $\alpha \leq \beta$, alors $D_\alpha(p||q) \leq D_\beta(p||q)$:

$$\underline{P_1}: 2^{(\alpha-1)D_\alpha(p||q)} = E_{X \sim p} \left[\left(\frac{p(X)}{q(X)} \right)^{\alpha-1} \right]$$

f est concave :
 $\mathbb{E}[f(x)] \leq f(\mathbb{E}[x])$

inégalité
de Jensen

$$\begin{aligned}
 &= \mathbb{E}_{x \sim p} \left[\left(\frac{p(x)}{q(x)} \right)^{\beta-1} \right]^{\frac{\alpha-1}{\beta-1}} \\
 &\leq \mathbb{E}_{x \sim p} \left[\left(\frac{p(x)}{q(x)} \right)^{\beta-1} \right]^{\frac{\alpha-1}{\beta-1}} \\
 &= \left(2^{(\beta-1) D_{\beta}(p||q)} \right)^{\frac{\alpha-1}{\beta-1}} \\
 &= 2^{(\alpha-1) D_{\beta}(p||q)}.
 \end{aligned}$$

$$\Rightarrow D_{\alpha}(p||q) \leq D_{\beta}(p||q).$$

$\Rightarrow$ Entropie conditionnelle est décroissante en α
 $\&$ $I_{\alpha}(X;Y)$ est croissante en α .

Lem.

$$\lim_{\alpha \rightarrow 1} D_{\alpha}(p||q) = D(p||q)$$

$$\lim_{\alpha \rightarrow 1} H_{\alpha}(X|Y) = H(X|Y)$$

$$\lim_{\alpha \rightarrow 1} I_{\alpha}(X;Y) = I(X;Y).$$

Compression
 $\uparrow$

- Max-entropie

- Min-entropie

Extraction
d'aléa

$$H_{\max}(X|Y) := H_{1/2}(X|Y)$$

$$H_{\min}(X|Y) := \lim_{\alpha \rightarrow \infty} H_{\alpha}(X|Y) = H_{\infty}(X|Y)$$

$$H_{\min}(X) = \min_{x \in \mathcal{X}} S(x)$$

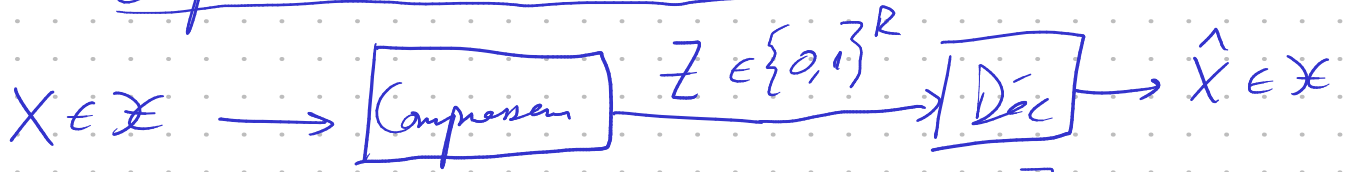
$$= -\log P_n[\text{deviner } x]$$

$$H_{\min}(X|Y) = -\log P_n[\text{devin } X \text{ sachant } Y].$$

- Min-information-mutuelle : $I_{\min}(X; Y) := I_{\frac{1}{2}}(X; Y)$.

↳ Transmission de données sur canaux bruités.

Compression de données one-shot



$$P_n[\text{erreur}] = P_n[\hat{X} \neq X].$$

Th: Il existe une paire (c, d) où $c: X \rightarrow \{0,1\}^R$ et $d \in \mathcal{P}(X | \{0,1\}^R)$,

t.g.
$$P_n[\hat{X} \neq X] \leq 2^{H_{\max}(X) - R}.$$

P_n : - On choisit $c(\cdot)$ au hasard selon une distribution t.g.

$$P_n[c(x) = c(x')] \leq 2^{-R} \quad \forall x \neq x'.$$

- On définit le décompression comme

$$d(\hat{x}|z) = \begin{cases} \frac{2^{-\frac{1}{2} S(\hat{x})}}{\sum_{x': c(x')=z} 2^{-\frac{1}{2} S(x')}} & \text{si } c(\hat{x})=z \\ 0 & \text{si } c(\hat{x}) \neq z \end{cases} = \frac{\sqrt{P_x(\hat{x})}}{\sum_{x': c(x')=z} \sqrt{P_x(x')}}.$$

Alas, la probabilité d'erreur est

$$P_n[\hat{X} \neq X | X=x] = \mathbb{E}_c \left[\frac{\sum_{\hat{x} \neq x, c(\hat{x})=c(x)} 2^{-\frac{1}{2}S(\hat{x})}}{\sum_{x': c(x')=c(x)} 2^{-\frac{1}{2}S(x')}} \right]$$

$$\leq \mathbb{E}_c \left[\frac{\sum_{\hat{x} \neq x, c(\hat{x})=c(x)} 2^{-\frac{1}{2}S(\hat{x})}}{2^{-\frac{1}{2}S(x)}} \right]$$

$$= 2^{\frac{1}{2}S(x)} \sum_{\hat{x} \neq x} \mathbb{P}_c[c(\hat{x})=c(x)] \cdot 2^{-\frac{1}{2}S(\hat{x})}$$

$$\leq 2^{\frac{1}{2}S(x)} \sum_{\hat{x} \neq x} 2^{-R - \frac{1}{2}S(\hat{x})}$$

$$= 2^{\frac{1}{2}S(x) - R} \sum_{\hat{x} \neq x} 2^{-\frac{1}{2}S(\hat{x})}$$

$$\leq 2^{\frac{1}{2}S(x) - R} \cdot \sum_{\hat{x}} \sqrt{p_x(\hat{x})}$$

$$= 2^{\frac{1}{2}S(x) + \frac{1}{2}H_{\max}(x) - R} \cdot 2^{\frac{1}{2}H_{\max}(x)}$$

$$P_n[\hat{X} \neq X] \leq \mathbb{E}_x \left[2^{\frac{1}{2}S(x) + \frac{1}{2}H_{\max}(x) - R} \right]$$

$$= 2^{\frac{1}{2}H_{\max}(x) - R} \cdot \underbrace{\mathbb{E}_x \left[2^{\frac{1}{2}S(x)} \right]}_{2^{\frac{1}{2}H_{\max}(x)}}$$

$$= 2^{H_{\max}(x) - R}$$

□

