Basic properties

$\log_2 \frac{1}{p_X(x)}$

$$H(X) = \sum_{x \in \mathcal{X}} p_X(x) \log_2 \frac{1}{p_X(x)}$$

: Average randomness in X

$$I(X;Y) = \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} p_{XY}(x,y) \log_2 \frac{p_{XY}(x,y)}{p_X(x)\, p_Y(y)}$$

: Amount of information Y gives about X

$$X \longrightarrow \boxed{p_{Y|X}} \longrightarrow Y$$

$$\frac{k}{n} \approx I(X;Y)$$

$$P_e \approx 0$$

$$I(Y;X) = I(X;Y)$$

$$k = n\, I(X;Y)$$

Conditional entropy

$H(X)$: Uncertainty / randomness in X

$X \rightarrow$ $p_X(\cdot) = 1$ deterministic

$H(X) = 0$ bits

$H(X)$ is max when X is uniform

all outcomes eq. likely

$(X, Y) \sim p_{XY}$ Conditional entropy of Y given X

$$H(Y|X) = \sum_{x \in \mathcal{X}} \underline{H(Y|X=x)}\, p_X(x)$$

$$= \sum_{x \in \mathcal{X}} \left[\sum_{y} \left(p_{Y|X}(y|x) \log_2 \frac{1}{p_{Y|X}(y|x)} \right) \right] p_X(x)$$

$$= \sum_{x} \sum_{y} p_{XY}(x,y) \log_2 \frac{1}{p_{Y|X}(y|x)}$$

$H(Y|X)$ is the avg amt of uncertainty left in Y after observing X.

$X = 0$

$$X \rightarrow \boxed{BS(p)} \rightarrow Y$$

$X \sim Ber(1/2)$

$p_{Y|X}(0|0) = 1-p$ $\qquad$ $p_{Y|X}(1|1) = 1-p$

$p_{Y|X}(0|1) = p$

$p_{Y|X}(1|0) = p$

$$H(X) = \sum_x p_X(x) \log \frac{1}{p_X(x)} = 0.5 \log \frac{1}{0.5} + 0.5 \log \frac{1}{0.5} = 1 \text{bit}$$

$$p_Y(0) = \sum_x p_{Y|X}(0|x)\, p_X(x)$$

$$p_Y(1) = \frac{1}{2}$$

$$= p(0|0)\, p_X(0) + p(0|1)\, p_X(1)$$

$$= (1-p)\, \frac{1}{2} + p \times \frac{1}{2}$$

$$Y \sim \mathrm{Ber}(1/2)$$

$$= \frac{1}{2}$$

$$\Rightarrow \quad H(Y) = 1 \text{ bit}$$

$$H(Y|X=0) = \sum_y p_{Y|X}(y|0) \log \frac{1}{p_{Y|X}(y|0)} = p(0|0) \log \frac{1}{p(0|0)} + p(1|0) \log \frac{1}{p(1|0)}$$

$$= H_2(p)$$

$$H(Y|X=1) = H_2(p)$$

$$H(Y|X) = \sum_x H(Y|X=x)\, p_X(x)$$

$$= H(Y|X=0)\, p_X(0) + H(Y|X=1)\, p_X(1)$$

$$= H_2(p)\, \frac{1}{2} + H_2(p)\, \frac{1}{2}$$

$$= H_2(p) \text{ bits.} \quad < 1 \quad \text{if } p < \frac{1}{2}$$

$$\text{or } p > \frac{1}{2}$$

$$H(Y|X) < H(Y)$$

Relation between entropy and mutual information

$$I(X;Y) = \sum_{x,y} p_{XY}(x,y) \log_2 \frac{p_{XY}(x,y)}{p_X(x)\, p_Y(y)}$$

$$= \sum_{x,y} p_{XY}(x,y) \left[\log_2 \frac{1}{p_Y(y)} + \log_2 p_{Y|X}(y|x) \right]$$

$$= \sum_{x,y} p_{XY}(x,y) \log_2 \frac{1}{p_Y(y)} + \sum_{x,y} p_{XY}(x,y) \log_2 p_{Y|X}(y|x)$$

$$= \sum_{y} \left(p_Y(y) \log_2 \frac{1}{p_Y(y)} \right) \underbrace{\left(\sum_{x} p_{X|Y}(x|y) \right)}_{=1} - \sum_{x,y} p_{XY}(x,y) \log_2 \frac{1}{p_{Y|X}(y|x)}$$

$$= H(Y) - H(Y|X)$$

$$I(X;Y) = H(Y) - H(Y|X)$$

$$= H(X) - H(X|Y)$$

Amt of info that X gives about Y = Avg Uncertainty about Y − Avg uncertainty left in Y after seeing X

Amt of info that Y gives about X = Avg unc abt X − "— X after seeing Y.

KL divergence

Kullback-Liebler divergence / Relative entropy

p_X , q_X pmf on $\mathcal{X}$

$$D(p_X \| q_X) = \sum_x p_X(x) \log \frac{p_X(x)}{q_X(x)}$$

$D(p_X \| p_X) = 0$ $D(p_X \| q_X) \geq 0$

$$I(X;Y) = \sum_{x,y} p_{XY}(x,y) \log_2 \frac{p_{XY}(x,y)}{p_X(x)\, p_Y(y)}$$

$$D(p \| q) = \sum_{x} p(x) \log_2 \frac{p(x)}{q(x)}$$

$$I(X;Y) = D(p_{XY} \| p_X p_Y)$$

p_{XY} defined on $\mathcal{X} \times \mathcal{Y}$

$p_X p_Y$ —"— $\mathcal{X} \times \mathcal{Y}$

$p_{XY}(x,y) \qquad q_{XY}(x,y) = p_X(x)\, p_Y(y)$

$D(p_{XY} \| q_{XY})$

$(X, Y) \sim p_{XY}$

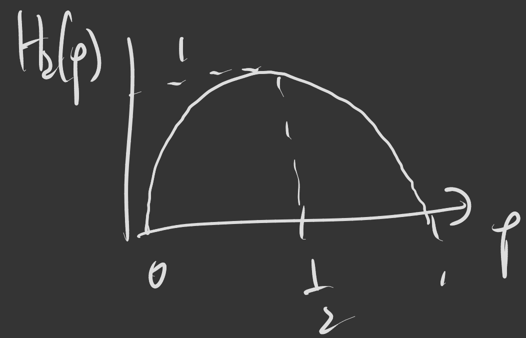
$(\tilde{X}, \tilde{Y}) \sim q_{XY} = p_X(x)\, p_Y(y)$

$I(X;Y) = 0 \Leftarrow X, Y$ are independent

$X \longrightarrow \boxed{BSC(p)} \longrightarrow Y$ $\quad p =$

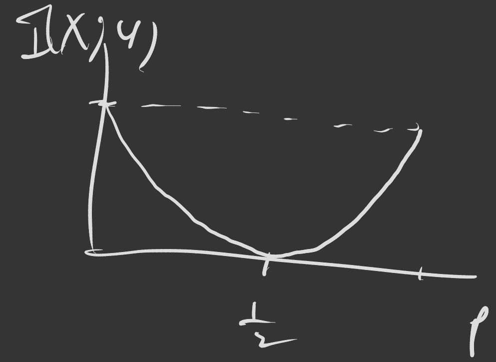
$\sim Ber(1/2)$

$H(Y|X) = H_2(p)$



$H(Y) = 1$ bit

$I(X;Y) = H(Y) - H(Y|X) = 1 - H_2(p)$



If $p = \frac{1}{2}$, X & Y indp

$p < \frac{1}{2}$, $p > \frac{1}{2}$ $\quad X$ & Y NOT INDEP

$$H(X) = \sum_{x} p_X(x) \log_2 \frac{1}{p_X(x)}$$

$$H(X/Y) = \sum_{x,y} p_{XY}(x,y) \log_2 \frac{1}{p_{X/Y}(x/y)}$$

$$I(X;Y) = \sum_{x,y} p_{XY}(x,y) \log_2 \frac{p_{XY}(x,y)}{p_X(x)\, p_Y(y)}$$

$$= H(X) - H(X|Y) = H(Y) - H(Y/X)$$

$$D(p\|q) = \sum_{x} p(x) \log_2 \frac{p(x)}{q(x)} \neq D(q\|p)$$

$$I(X;Y) = D(p_{XY} \| p_X p_Y)$$

$$R = \frac{\mathbb{E} L}{n} \approx H(p_X) + D(p_X \| q_X)$$

$X^n \sim p_X$ Design based on q_X

$X \sim \mathrm{Ber}(p)$ $\longrightarrow$

$X' \sim \mathrm{Ber}(q)$

$\{0,1\} \rightarrow$ Support/ alphabet.

$$p_X(x) \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$$

$$q_X(x) = \begin{cases} q & x = 1 \\ 1-q & x = 0 \end{cases}$$

Chain rule of entropy

$$(X, Y) \sim p_{XY}$$

$$H(X,Y) = \sum_{x,y} p_{XY}(x,y) \log_2 \frac{1}{p_{XY}(x,y)}$$

$$= \sum_{x,y} p_{XY}(x,y) \log \frac{1}{p_X(x)\, p_{Y|X}(y|x)}$$

$$= \sum_{x,y} p_{XY}(x,y) \left[\log_2 \frac{1}{p_X(x)} + \log_2 \frac{1}{p_{Y|X}(y|x)} \right]$$

$$= \sum_{x,y} p_{XY}(x,y) \log_2 \frac{1}{p_X(x)} + \sum_{x,y} p_{XY}(x,y) \log_2 \frac{1}{p_{Y|X}(y|x)}$$

$$= \sum_{x} \left[\left(p_X(x) \log_2 \frac{1}{p_X(x)} \right) \left(\sum_{y} p_{Y|X}(y|x) \right) \right] + H(Y|X)$$

$$H(x,y) = H(x) + H(y/x)$$

$$\sum_y p(y|x) = 1$$

$$H(X,Y) = H(Y,X) = H(Y) + H(X/Y)$$

$$H(X_1\ X_2\ X_3) = H(X_1) + H(X_2|X_1) + H(X_3|X_1 X_2)$$

OPTIMAL JOINT

SEQUENTIAL (OPTIMAL)

$$\frac{H(X_1^n X_2^n X_3^n)}{n}$$

$$H(X_1 \cdots X_n) = H(X_1) + H(X_2 | X_1) + H(X_3 | X_2 X_1) + \cdots + H(X_n | X_{n-1} \cdots X_1)$$

$$H(X^n) = \sum_{i=1}^{n} H(X_i | X_1 \cdots X_{i-1})$$

$$X_1 X_2 \cdots X_n$$

$X_1 \cdots X_n \sim$ indep. $\qquad P_{X_i | X_1 \cdots X_{i-1}} = P_{X_i}$

$$H(X_i | X_1 \cdots X_{i-1}) = H(X_i)$$

$$H(X^n) = H(X_1) + H(X_2) \cdots + H(X_n)$$

Chain rule of mutual information

$$I(X;YZ) = \sum_{x\,y\,z} p_{XYZ}(x\,y\,z) \log_2 \frac{p_{XYZ}(x,y,z)}{p_X(x) \cdot p_{YZ}(y,z)}$$

$$= \sum_{x,y,z} p_{XYZ}(x,y,z) \log_2 \frac{p_{X|YZ}(x|y,z)}{p_X(x)}$$

$$= \sum_{x,y,z} p_{XYZ}(x,y,z) \log_2 \left(\frac{p_{XY}(x,y) \cdot p_{Z|XY}(z|x,y)}{p_X(x)\, p_Y(y) \cdot p_{Z|Y}(z|y)} \right)$$

$$= \sum_{x,y,z} p(x,y,z) \left[\log \left(\frac{p(x,y)}{p(x)\, p(y)} \right) + \log \left(\frac{p_{Z|XY}(z|x,y)}{p(z|y)} \right) \right.$$

$$= \sum_{x,y,z} p(x,y,z) \log_2 \frac{p(x,y)}{p(x)\,p(y)} + \sum_{x,y,z} p(x,y,z) \times \log\left(\frac{p(z|x,y)}{p(z|y)} \right.$$

$$= \sum_{x,y} \log \frac{p(x,y)}{p(x)p(y)} \sum_{z} p(x,y,z)$$

$$+ \sum_{x,y,z} p(x,y,z) \log_2 \left(\frac{p(x|y)\,p(z|x,y)}{p(x|y)\,p(z|y)} \right)$$

$$= \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)\,p(y)} + \sum_{x,y,z} p(x,y,z) \log\left(\frac{p(x,z|y)}{p(x|y)\,p(z|y)} \right.$$

$$I(X;YZ) = I(X;Y) + I(X;Z|Y)$$

$$I(X;YZ) = H(YZ) - H(YZ|X)$$

Chain rule of entropy

$$= H(Y) + H(Z|Y) - \left(H(Y|X) + H(Z|XY) \right)$$

$$= \underbrace{H(Y) - H(Y|X)} + \underbrace{H(Z|Y) - H(Z|X,Y)}$$

$$= I(X;Y) + I(X;Z|Y)$$

$$I(XY;UV|W) = H(XY|W) - H(XY|UVW)$$

$$= H(UV|W) - H(UV|XYW)$$

$$H(AB|C) = H(A|C) + H(B|AC)$$

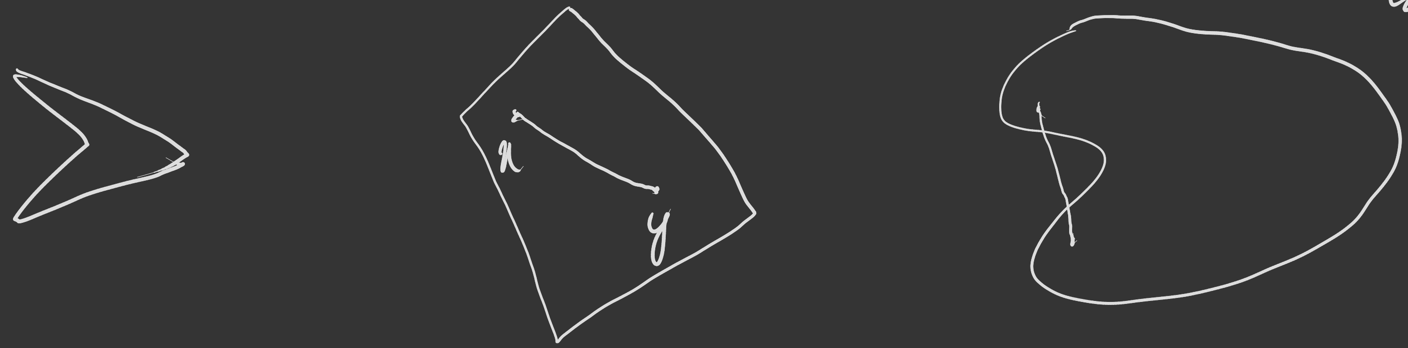
$$I(X;Y_1 \ldots Y_n \mid W) = \sum_{i=1}^{n} I(X;Y_i \mid Y_1 \ldots Y_{i-1}, W)$$

$$= I(X;Y_1 \mid W) + I(X;Y_2 \mid Y_1 W) + I(X;Y_3 \mid \underline{Y_1 Y_2 W})$$

$$+ \ldots$$

Convex sets

NOT CONVEX

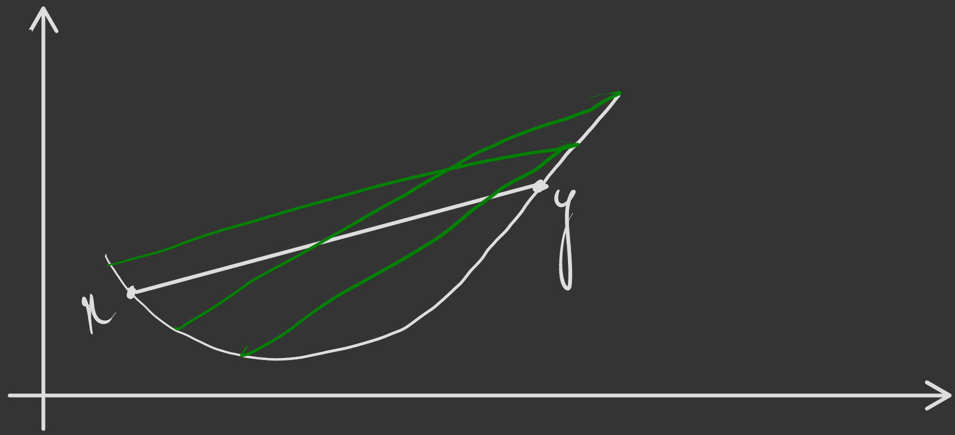


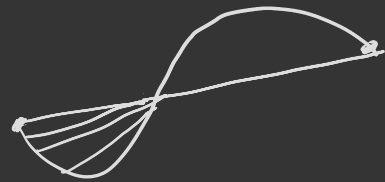
$A \subseteq \mathbb{R}^n$ is convex if $\forall\ \alpha \in [0,1]$, $x, y \in A$

$$\alpha x + (1-\alpha) y \in A$$

$\alpha = 0 \qquad \alpha x + (1-\alpha) y = y$

$\alpha = 1 \qquad$ — " — $=$ — $= x$

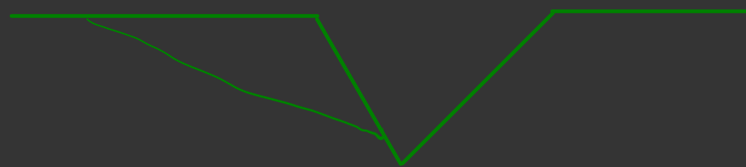




Convex functions: $f: \mathbb{R} \to \mathbb{R}$ is convex in $[a,b]$ if

$$f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y) \qquad \forall x, y \in [a,b]$$

Strictly convex :



$f: \mathbb{R} \to \mathbb{R}$ is concave if $-f$ is convex.

Theorem: $f: \mathbb{R} \to \mathbb{R}$ is convex in $[a, b]$ if & only if (assuming f'' exists)

$$f''(x) \geq 0 \quad \forall x \in [a, b]$$

Strictly convex: > 0

Are these convex:

① $\log x$

② x^2

(3) $x \log x$

Jensen's inequality

If f is a convex function over $\mathbb{R}$ & X any rv,

$$\mathbb{E} f(X) \geq f(\mathbb{E} X)$$

① $f(x) = x^2$

② $f(x) = \log_2 x$

Important consequence: the log-sum inequality

Let $(\alpha_1 \ldots \alpha_m)$ & $(\beta_1 \ldots \beta_m)$ be non-negative real nos

st $\beta_i > 0 \Rightarrow \alpha_i > 0$

Then,

$$\sum_{i=1}^{m} \alpha_i \log_2 \frac{\alpha_i}{\beta_i} \geq \left(\sum_{i=1}^{m} \alpha_i\right) \log_2 \frac{\left(\sum_{i=1}^{m} \alpha_i\right)}{\left(\sum_{i=1}^{m} \beta_i\right)}$$

Proof: $y \log y$ is convex.

$$D(p \| q) = \sum_{x} p(x) \log_2 \frac{p(x)}{q(x)}$$

Apply log-sum inquality