$X$ v.a. , $X \in \mathcal{X} = \{x_1, \dots, x_K\}$
d.d.p. $p = (p_1, \dots, p_K) \in \mathbb{R}^K$

$$H(X) = -\sum_{i=1}^{K} p_i \log p_i = \mathop{\mathbb{E}}_{X \sim p_X}[-\log p(X)]$$

"$X \sim p$" : $X$ ha d.d.p $p$

$$D(p_{XY} \| p_X \cdot p_Y) \overset{(def)}{=} I(X;Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)\, p(y)} =$$

$$= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)\, p(y)} = \mathop{\mathbb{E}}_{(X,Y) \sim p_{XY}} \left[ \log \frac{p(X,Y)}{p(X) \cdot p(Y)} \right]$$

$$D(p \| q) = \sum_{i=1}^{K} p_i \log \frac{p_i}{q_i} = \mathop{\mathbb{E}}_{X \sim p_X} \left[ \log \frac{p(X)}{q(X)} \right]$$

<u>Entropia congiunta</u> di 2 v.a. $X, Y$ $\qquad X \in \mathcal{X}, Y \in \mathcal{Y} \to (X,Y) \in \mathcal{X} \times \mathcal{Y}$

$p_{XY}$ d.d.p congiunta

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

$$= \mathbb{E}_{(X,Y) \sim p_{XY}} [-\log p(X,Y)] = \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} p(x,y) [-\log p(x,y)] =$$

$$= -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log p(x,y)$$

<u>Entropia condizionata</u> di $Y$ data $X$

Probabilità condizionata di $Y=y$ data $X=x$:

$$\Pr[Y=y \mid X=x] = \frac{\Pr[Y=y, X=x]}{\Pr[X=x]}$$

$$\boxed{p(y|x) = \frac{p(x,y)}{p(x)}}$$

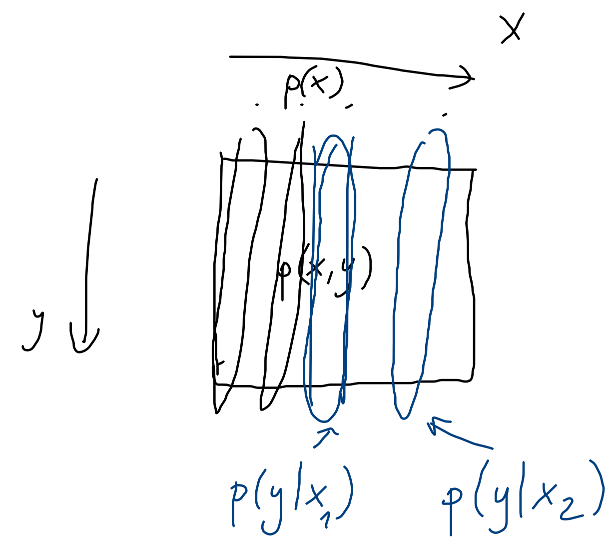
$$\to \Pr[Y=y, X=x] = \Pr[X=x] \cdot \Pr[Y=y \mid X=x]$$

$$\Pr[X=x, Y=y] = \Pr[Y=y] \cdot \Pr[X=x \mid Y=y]$$

Formula di Bayes:

$$\Pr[Y=y \mid X=x] = \frac{\Pr[Y=y] \cdot \Pr[X=x \mid Y=y]}{\Pr[X=x]}$$

Entropia condizionata di Y data X :




$$\boxed{H(Y|X)} = \mathbb{E}_{(X,Y)\sim p_{XY}}\left[-\log p(Y|X)\right] =$$

$$= \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p(x,y)\left[-\log p(y|x)\right] =$$

$$= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \overbrace{\boxed{p(x,y)}}^{=\,p(x)\cdot p(y|x)} \log p(y|x) = -\sum_x \sum_y p(x)\, p(y|x) \log p(y|x)$$

$$= \sum_{x\in\mathcal{X}} p(x)\Big(-\overbrace{\boxed{\sum_{y\in\mathcal{Y}} p(y|x)\log p(y|x)}}^{H(Y|X=x)}\Big)$$

$$= \sum_{x\in\mathcal{X}} \underbrace{p(x)}_{\geq 0} \underbrace{H(Y|X=x)}_{\geq 0} \geq 0$$

← è $=0$ se e solo se la Y è completamente determinata dalla X

Per ogni v.a. $X$,

$$H(X) \geq 0 \quad \text{sempre}$$

Ma quanto grande può essere $H(X)$?

Prop. Si ha sempre $H(X) \leq \log |\mathcal{X}|$ e si ha uguaglianza se e solo se la v.a. $X$ è uniformemente distribuita su $\mathcal{X}$.

Dim. Consideriamo la d.d.p. $p_X$ della v.a. $X$

e la d.d.p. uniforme $\mathcal{U}$ sull'alfabeto $\mathcal{X} = \{x_1, \dots, x_K\}$

$\rightarrow \mathcal{U} = (1/K, 1/K, \dots, 1/K)$

(Gibbs)

$$0 \leq D(p_X \| \mathcal{U}) = \sum_{i=1}^{K} p_i \log \frac{p_i}{1/K} = \sum_{i=1}^{K} p_i \log (p_i K) =$$

$$= \underbrace{\sum_{i=1}^{K} p_i \log p_i}_{-H(X)} + \underbrace{\sum_{i=1}^{K} p_i}_{1} (\log K) = -H(X) + \log K$$

$$\Rightarrow H(X) \leq \log K = \log |\mathcal{X}|$$

Prop. $I(X;Y) \stackrel{?}{=} H(X) - H(X|Y)$

$p(x,y) = p(y)\cdot p(x|y)$

Dim. $I(X;Y) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)\,p(y)}$

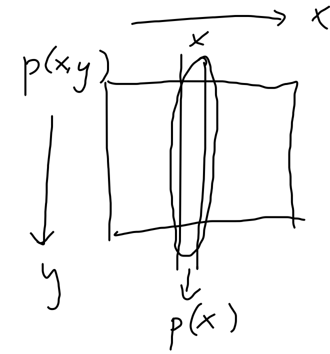
$$= \sum_x \sum_y p(x,y) \log \frac{p(x|y)}{p(x)} = \sum_x \sum_y p(x,y) \left[\log p(x|y) - \log p(x)\right]$$

non dipende da y

$$= \sum_x \sum_y p(x,y) \log p(x|y) - \sum_x \underbrace{\sum_y p(x,y)}_{p(x)} \log p(x)$$

$H(X|Y=y)$

$$= \sum_x \sum_y \underbrace{p(x,y)}_{p(y)\cdot p(x|y)} \log p(x|y) \underbrace{- \sum_x p(x) \log p(x)}_{+H(X)}$$




$$= \sum_y \sum_x p(y)p(x|y) \log p(x|y) \qquad + H(X)$$

$$= -\sum_y p(y)\, H(X|Y=y) \qquad + H(X) = -H(X|Y) + H(X).$$

QED

Poiché $I(Y;X) = I(X;Y)$, ho anche

$$0 \le I(X;Y) = H(Y) - H(Y|X)$$

$$0 \le I(X;Y) = H(X) - H(X|Y)$$

Poiché $I(X;Y) \geq 0$, ho $H(Y) \geq H(Y|X)$, con uguaglianza sse $X$ e $Y$ sono v.a indipendenti

e $H(X) \geq H(X|Y)$, " " " " "

Poiché $H(Y|X) \geq 0$, ho $I(X;Y) \le H(Y)$, con uguaglianza sse la $Y$ è completamente determinata dalla $X$

$H(Y|X) = 0$ se e solo se la $Y$ è completamente determinata dalla $X$

$(Y = f(X))$

# Regola della catena

$p(x,y) = p(x)\, p(y|x)$

Prop. $H(X,Y) = H(X) + H(Y|X)$

Dim. $H(X,Y) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log \boxed{p(x,y)}$

$= -\sum_x \sum_y p(x,y) \left[ \log p(x) + \log p(y|x) \right] =$

$= -\sum_x \underbrace{\sum_y p(x,y)}_{p(x)} \log p(x) \underbrace{- \sum_x \sum_y p(x,y) \log p(y|x)}_{+H(Y|X)}$

$= H(X) + H(Y|X)$ . QED

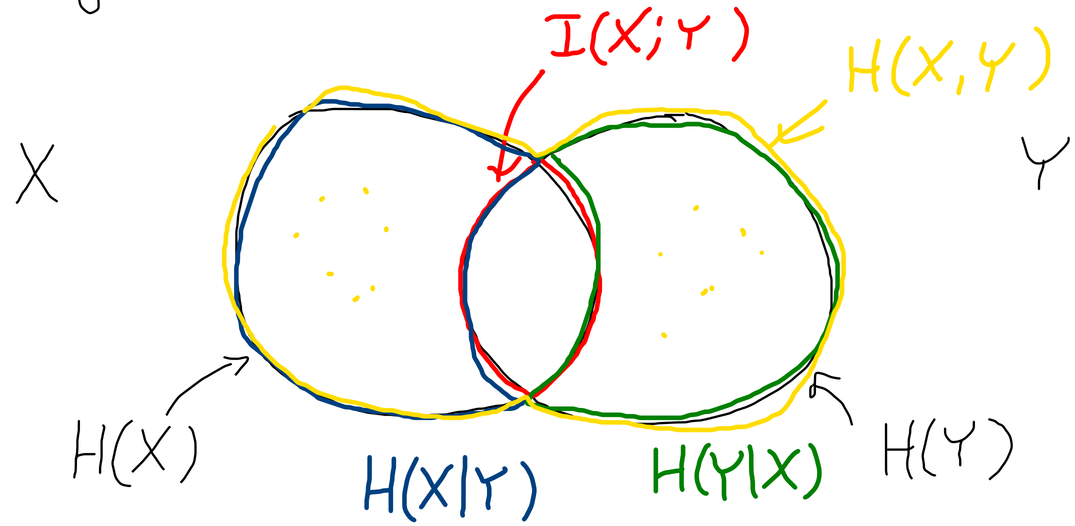
Per simmetria, $\boxed{H(X,Y) = H(Y) + H(X|Y)}$

→ Corollario. $I(X;Y) = H(X) + H(Y) - H(X,Y)$

Dim. $I(X;Y) = H(X) - H(X|Y) = H(X) - \left[ H(X,Y) - H(Y) \right] = H(X) + H(Y) - H(X,Y)$

QED.

# Diagrammi informazionali




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

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

$\cap$ intersezione

differenza insiemistica

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$