Simboli : $0, 1,$ $a, b, c, \dots$

Un Alfabeto è un insieme finito di simboli

$\mathcal{X} = \{0, 1\}$ $\mathcal{X} = \{a, b, c, \dots, z\}$ $\mathcal{X} = \{0, 1, \dots, 255\}$

Una parola (o stringa) su un alfabeto $\mathcal{X}$ è una sequenza di simboli di $\mathcal{X}$

Es. 010010 shannon 13 102 207

$\mathcal{X} = \{x_1, x_2, \dots, x_K\}$

Distribuzione di probabilità (d.d.p.)

$$P_X = \{p_X(x), \ x \in \mathcal{X}\}$$

dove $p_X(x) \geq 0$ e $\sum_{x \in \mathcal{X}} p_X(x) = 1$

$$\sum_{i=1}^{K} p_X(x_i) = 1$$

Variabile aleatoria (v.a.)

$$X, Y \in \mathcal{X}$$

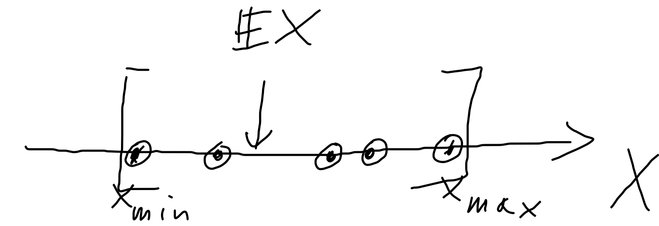
$$P_X \rightsquigarrow (p_X(x_1), p_X(x_2), \dots, p_X(x_K)) \in \mathbb{R}^K$$

$$p = (p_1, p_2, \dots, p_K) \in \mathbb{R}^K$$

$$P_Y \rightsquigarrow (p_Y(x_1), p_Y(x_2), \dots, p_Y(x_K))$$

$$q = (q_1, q_2, \dots, q_K)$$

$$P_{XY} = \{ \overset{(\geq 0)}{p_{XY}(x,y)} \ , \ x \in \mathcal{X}, y \in \mathcal{Y} \}$$



$$\sum_{(x,y)\in \mathcal{X}\times\mathcal{Y}} p_{XY}(x,y) = 1$$

Es. $\mathcal{X} = \mathcal{Y} = \{0, 1\}$

X \ Y	0	1	
0	1/4	1/12	1/3
1	1/2	1/6	2/3

$$1/4 + 1/12 + 1/2 + 1/6 = 1$$

$$\Pr[X=0] = 1/4 + 1/12 = 1/3$$

$$\Pr[X=1] = 1/2 + 1/6 = 2/3$$

$$P_X = (1/3, 2/3)$$

$$P_Y = (3/4, 1/4)$$

Per una v.a. $X \in \mathbb{R}$ (discreta), il valore atteso è $\mathbb{E}[X] = \mathbb{E}X = \sum_{i=1}^{k} x_i \cdot P_X(x_i) = \sum_{i=1}^{k} p_i x_i$

Divergenza informazionale

Mutua informazione

Entropia

$P = (p_1, \dots, p_k) \in \mathbb{R}^k$ $\quad$ $Q = (q_1, \dots, q_k) \in \mathbb{R}^k$

$$D(p \| q) \stackrel{def}{=} \sum_{i=1}^{k} p_i \log \frac{p_i}{q_i}$$

Disuguaglianza di Jensen

È una sorta di "distanza":

1. È continua nei suoi argomenti
2. $D(p\|q) \geq 0$ per ogni d.d.p. p e q (disuguaglianza di Gibbs)
3. $D(p\|q) = 0$ se e solo se $p = q$

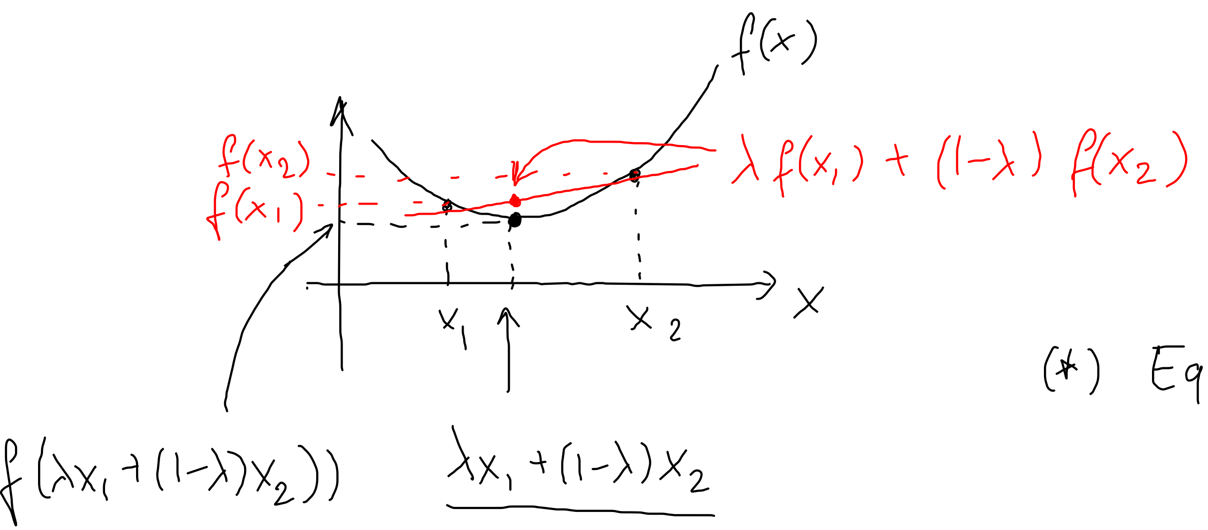
Disuguaglianza di Jensen. Se $f : \mathbb{R} \to \mathbb{R}$ è una funzione convessa e X è una v.a. (discreta) allora

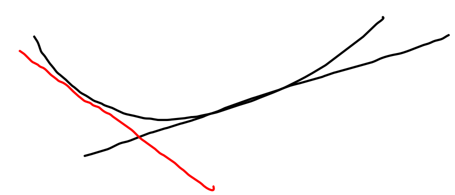
$$\mathbb{E} f(X) \geq f(\mathbb{E} X) \quad \text{per ogni v.a. } X. \qquad (*)$$

Una funzione $f : \mathbb{R} \to \mathbb{R}$ è <u>convessa</u> nell'intervallo <u>(a,b)</u> se

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

per ogni $x_1, x_2 \in (a,b)$ e $\lambda \in [0,1]$





(*) Equivalentemente,

$$\sum_{i=1}^{k} p_i \, f(x_i) \geq f\left(\sum_{i=1}^{k} p_i \, x_i\right).$$