# Teoria dell'Informazione

Simone Alessandro Casciaro

08 Novembre 2024

## Lezione 10: Canale

Ricordando quanto detto in lezioni precedenti, quando spediamo un messaggio sul canale usiamo due variabili aleatorie: $X$, che estrae un simbolo dalla Sorgente $< \mathbb{X}, \mathbb{P} >$ e $Y$, che indica invece il simbolo ricevuto dal Ricevente.

$$X \longrightarrow Y$$

Definiamo una funzione $g : \mathbb{Y} \rightarrow \mathbb{X}$ con $g(y) = x$
La funzione $g$ cerca di associare il simbolo $y$ ricevuto dal canale al simbolo $x$ che è stato originariamente spedito.
La funzione $g$ potrebbe sbagliare, se sul canale è presente del rumore.
Chiamiamo $p_e$ la probabilità che $g$ sbagli.

$$p_e = \mathbb{P}\left(g(y) \neq x\right)$$

## Teorema: Disuguaglianza di Fano

### Ipotesi

Siano $X, Y$ due variabili aleatorie con valori in $\mathbb{X}, \mathbb{Y}$ entrambi finiti
Sia $g : \mathbb{Y} \rightarrow \mathbb{X}$ la funzione definita sopra.
Sia $p_e$ la probabilità d'errore. $p_e = \mathbb{P}\left(g(y) \neq x\right)$

### Tesi

$$p_e \geq \frac{H(X|Y) - 1}{\log |X|}$$

## Dimostrazione

Sia $E$ una variabile aleatoria bernoulliana definita come segue.

$$E = \begin{cases} 1 & g(y) \neq x \\ 0 & g(y) = x \end{cases}$$

Ricordando la Chain Rule per l'entropia,

$$H(E, X|Y) = H(E|Y) + H(X|E, Y) = H(X|Y) + H(E|X, Y)$$

Analizziamo ogni elemento uno per volta.

- $H(E|Y) \leq 1$ perché $E$ è una variabile aleatoria bernoulliana, come visto nella Lezione 4
- $H(E|X, Y) = 0$ perché una volta che conosciamo $X$ e $Y$, sappiamo già se $E = 1$ o $E = 0$, dunque non abbiamo informazione aggiuntiva spedendo $E$.
- $H(X|E, Y)$
  $$\begin{aligned} H(X|E, Y) &= \sum p_E H(X|E = e, Y) \\ &= \mathbb{P}(E = 0) H(X|E = 0, Y) + \mathbb{P}(E = 1) H(X|E = 1, Y) \\ &= (1 - p_e) * 0 + p_e H(X|E = 1, Y) \\ &= p_e H(X|E = 1, Y) \\ &\leq \log(|\mathbb{X}| - 1) \end{aligned}$$
  In realtà, per semplicità nei calcoli, Fano non considera il $-1$ nel logaritmo. Volendo essere precisi può essere messo, perché sapendo che $E = 1$ (e quindi sapendo che la funzione $g$ ha predetto un dato errato) si può escludere uno degli elementi della sorgente $\mathbb{X}$ (ovvero quello corretto).

Mettendo insieme i pezzi, si ottiene: $p_e \log(|\mathbb{X}| - 1) + 1 \geq H(X|Y)$
E dunque:

$$p_e \geq \frac{H(X|Y) - 1}{\log(|\mathbb{X}| - 1)}$$

# Canale

Definiamo un Canale come la tripla $< \mathbb{X}, \mathbb{Y}, \mathbb{P} >$ dove $\mathbb{X}$ è l'insieme dei simboli emessi dalla sorgente, $\mathbb{Y}$ è l'insieme dei simboli ricevuti dal ricevente e $\mathbb{P}$ è la matrice stocastica di canale che indica tutte le probabilità condizionate $p(y|x) = \mathbb{P}(Y = y|X = x)$
Definiamo un messaggio spedito dalla sorgente $x^n = (x_1, \dots, x_n)$
e Definiamo un messaggio ricevuto dalla sorgente $y^n = (y_1, \dots, y_n)$
In generale, si ha che $p(a, b, c) = p(a|b, c) * p(b|c) * p(c)$ (Nota: questo vale anche per $n$ simboli)

# Canale Discreto Senza Memoria

Definiamo un Canale Discreto Senza Memoria un canale che non viene influenzato da ciò che è stato mandato in precedenza e da ciò che verrà mandato in futuro.
Dunque, se $a, b, c \in \mathbb{X}$, allora $p(a, b, c) = p(a) * p(b) * p(c)$

Vogliamo calcolare $p(y^n|x^n)$

$$\begin{aligned} p(y^n|x^n) &= p(y_n|y^{n-1}x^n) * p(y_{n-1}|y^{n-2}x^n) * \dots * p(y_3|y^2x^n) * p(y_2|y_1x^n) * p(y_1|x^n) \\ &= p(y_n|x_n) * p(y_{n-1}|x_{n-1}) * \dots * p(y_3|x_3) * p(y_2|x_2) * p(y_1|x_1) \\ &= \prod_{i=1}^{n} p(y_i|x_i) \end{aligned}$$

# Capacità del Canale

La Capacità di un Canale $< \mathbb{X}, \mathbb{Y}, \mathbb{P} >$ è definita come

$$C = \max_{p(x)} I(X, Y)$$

Ricordiamo:

$$0 \leq I(X, Y) = \begin{cases} H(X) - H(X|Y) \leq \log |X| \\ H(Y) - H(Y|X) \leq \log |Y| \end{cases}$$

Quindi

$$0 \leq C \leq \min\{\log|X|, \log|Y|\}$$

# Esercizio

Si supponga di disporre di un canale binario senza rumore. Calcolare $C$
Usando la relazione precedente, sappiamo che $C \leq \min\{\log|X|, \log|Y|\} = 1$

Inoltre, $I(X,Y) = H(X) - H(X|Y)$ ma dato che il canale è senza rumore, $H(X|Y) = 0$
Dato che vogliamo massimizzare l'entropia, sappiamo che l'entropia di una variabile aleatoria bernoulliana è massima quando $p = \frac{1}{2}$ e, in quel caso, $H(X) = 1$.

Dunque, $\boxed{C = 1}$