L'entropia media per simbolo di sorgente è:

$\frac{H(U^K)}{K}$; se $U_1, U_2, \dots, U_K$ sono indipendenti e identicamente distribuiti (sorgente stazionaria e senza memoria)

allora $H(U^K) = K \cdot H(U_i) \quad \forall i$

→ in questo caso l'entropia media per simbolo di sorgente è $H(U_1) (= H(U))$.

$H(U)$

L'ideale sarebbe ottenere $V^K = (V_1, \dots, V_K) = (U_1, \dots, U_K) = U^K$

Se invece $U_i \neq V_i$ ho un errore nelle posizione i-esima. → $P_{e/i} \triangleq Pr[U_i \neq V_i]$

Probabilità di errore media: $P_e \triangleq 1/K \sum_{i=1}^{K} P_{e/i}$

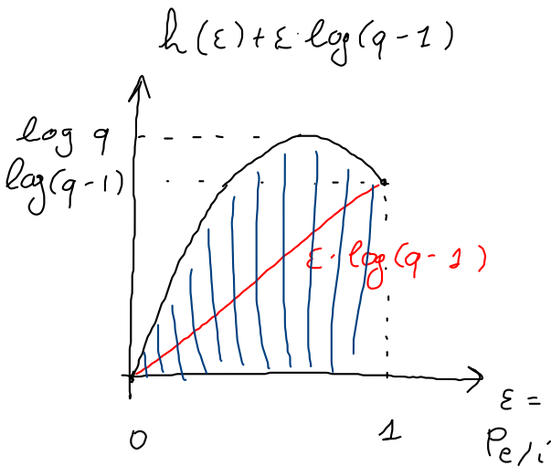
Numero di errori attesi su una sequenza lunga K è $K \cdot P_e$

Vogliamo dimostrare che se R (tasso) > C (capacità), allora $P_e \geq$ costante positiva > 0

$P_{e/i} = Pr[U_i \neq V_i]$

Useremo la disuguaglianza di Fano:

> Se $X, \hat{X}$ sono v.a. e $Pr[\hat{X} \neq X] = \varepsilon$, e $X \in \mathcal{X}$ con $|\mathcal{X}| = K$
> allora $H(X|\hat{X}) \leq h(\varepsilon) + \varepsilon \cdot \log(K-1)$.

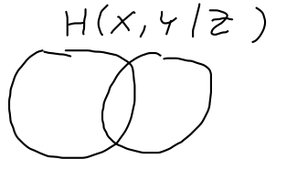


Applichiamola con $X = U_i$, $\hat{X} = V_i$, $\varepsilon = Pr[U_i \neq V_i] = P_{e/i}$, $K = |\mathcal{U}| = q$:

$\Rightarrow H(U_i|V_i) \leq h(P_{e/i}) + P_{e/i} \cdot \log(q-1)$. (per ogni $i = 1, \dots, K$)

Per U^K e V^K ho (per la regola della catena)

$$H(U^K|V^K) = H(U_1|V^K) + H(U_2|U_1, V^K) + H(U_3|U_1, U_2, V^K) + \dots$$
$$\leq \sum_{i=1}^{K} H(U_i|V_i) \leq \sum_{i=1}^{K} \left[h(P_{e/i}) + P_{e/i} \log(q-1) \right]$$

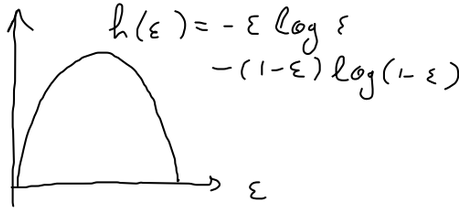


$= H(X|Z) + H(Y|X,Z)$

$$\frac{1}{K} H(U^K|V^K) \leq \underbrace{\frac{1}{K} \sum_{i=1}^{K} h(P_{e/i})}_{\text{Preferirei } h(P_e)} + \underbrace{\left(\frac{1}{K} \sum_{i=1}^{K} P_{e/i} \right)}_{P_e} \log(q-1)$$

$$\frac{1}{K} H(U^K|V^K) \leq \underbrace{\frac{1}{K}\sum_{i=1}^{K} h(P_{e/i})}_{\text{Preferirei } h(P_e)} + \underbrace{\left(\frac{1}{K}\sum_{i=1}^{K} P_{e/i}\right)}_{P_e} \log(q-1)$$

(ma h non è lineare...)



h è concava
($h'' < 0$)

Disuguaglianza di Jensen:

Se f è convessa, e Z è una v.a. qualunque,

allora $E[f(Z)] \geq f(E[Z])$

→ Se f è concava → $-f$ è convessa → $E[-f(Z)] \geq -f(E[Z])$

$-E[f(Z)] \geq -f(E[Z]) \rightarrow E[f(Z)] \leq f(E[Z])$

⇓

Considero $Z \in \{P_{e/1}, P_{e/2}, \dots, P_{e/K}\}$ con d.p. $= (1/K, 1/K, \dots, 1/K)$

$\overset{\text{Jensen}}{\Longrightarrow}$ (a h e a Z)

$E[h(Z)] \leq h(E[Z])$

$\| \qquad \| \qquad (P_e)$

$1/K \sum_{i=1}^{K} h(P_{e/i}) \qquad h\left(\frac{1}{K}\sum_{i=1}^{K} P_{e/i}\right)$

$\Rightarrow 1/K \sum_{i=1}^{K} h(P_{e/i}) \leq h(P_e)$

$\Rightarrow 1/K\, H(U^K|V^K) \leq h(P_e) + P_e \log(q-1)$

$$\frac{1}{k} H(U^k|V^k) \leq h(P_e) + P_e \log(q-1)$$

$$\frac{1}{k}\left[H(U^k) - I(U^k;V^k)\right]$$

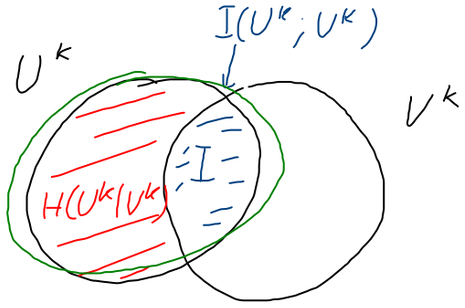
$$\frac{1}{k} H(U^k) - \frac{I(U^k;V^k)}{k}$$

$H(U)$

(Per il lemma 4.1)

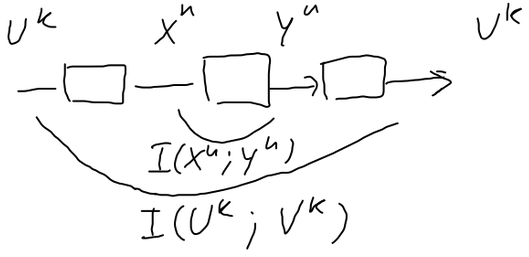
$$\geq H(U) - \frac{nC}{k}\cdot 1/R$$

$$R = k/n$$



Per il 2° teorema di elaborazione dati

$$\geq H(U) - \frac{I(X^n;Y^n)}{k}$$

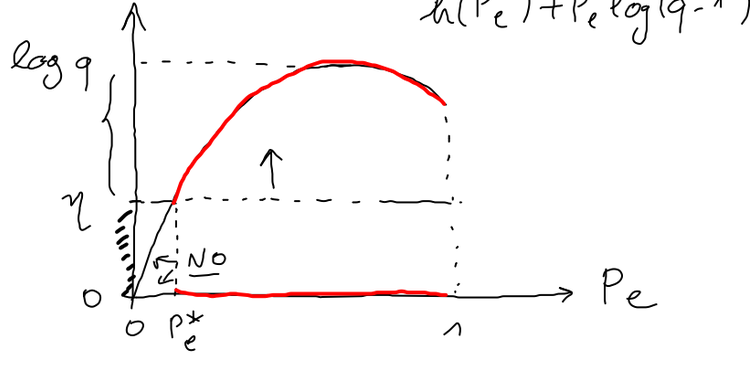


$$I(U^k;V^k) \leq I(X^n;Y^n)$$

Teorema inverso di Fano (T4.1): Se $H(U) > C/R$ allora la probabilità di errore rimane limitata inferiormente da una costante positiva.

Disequazione: $h(P_e) + P_e \log(q-1) \geq H(U) - C/R = \eta > 0$

$$\Rightarrow P_e \geq P_e^* > 0$$



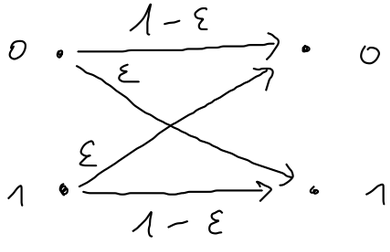
Se $q=2$ e $H(U)=1$, $\quad R > C$

allora il teorema dice che se $1 > C/R$

allora la prob. di errore rimane limitata inferiormente da una costante positiva.

Parte diretta del teorema di codifica di canale

Diamo una dimostrazione rigorosa solo per il canale binario simmetrico. (CSB(ε))



Capacità $= 1 - h(\varepsilon)$

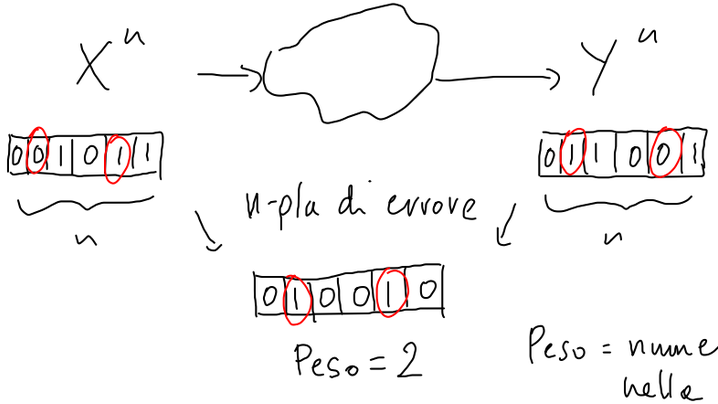
Assumiamo $\varepsilon < 1/2$.

- Codice con M parole di codice $\{x_1, \dots, x_M\}$ (come costruirlo?)
  ↘ tutte equiprobabili (la sorgente ha d.p. uniforme)
- Decodifica a massima verosimiglianza

$P_e(x_i) \triangleq$ probabilità di errore di decodifica quando viene trasmessa x_i

$P_e = \frac{1}{M} \sum_{i=1}^{M} P_e(x_i)$ probabilità di errore media

Dimostreremo che per ogni $\delta > 0$ e n sufficientemente grande, esiste un codice con tasso $\approx C$ e $P_e < \delta$.



Peso = numero di bit pari a 1 nella n-pla di errore

Probabilità $= \varepsilon^2 (1-\varepsilon)^4$ $\quad (1-\varepsilon)\,\varepsilon\,(1-\varepsilon)(1-\varepsilon)\,\varepsilon\,(1-\varepsilon)$

$\rightarrow$ Se la n-pla di errore ha peso w, la sua probabilità è $\varepsilon^w (1-\varepsilon)^{n-w}$

W = "Numero di bit errati ricevuti" è una v.a. con valore atteso $n \cdot \varepsilon$

$W = \sum_{i=1}^{n} Z_i$, $Z_i = \begin{cases} 1 \text{ se ho errore in posizione } i \\ 0 \text{ altrimenti} \end{cases}$ $\rightarrow$ $\Pr[Z_i = 1] = \varepsilon$, $\Pr[Z_i = 0] = 1 - \varepsilon$

$\rightarrow \mathrm{Var}(Z_i) = \varepsilon(1-\varepsilon)$

e varianza $\mathrm{Var}\left(\sum_{i=1}^{n} Z_i\right) = \sum_{i=1}^{n} \mathrm{Var}(Z_i)$ (Le Z_i sono tra loro indipendenti)

$= n\,\varepsilon(1-\varepsilon)$.

Per la disuguaglianza di Chebyshev:

$$\Pr[\, w > E[w] + \alpha \,] \leq \frac{\mathrm{Var}(w)}{\alpha^2}$$

Applicando con $\alpha = \sqrt{\frac{n\,\varepsilon(1-\varepsilon)}{\delta/2}}$

$$\Pr[w > E[w] + \alpha] \leq \frac{\cancel{n\,\varepsilon(1-\varepsilon)}}{\cancel{n\,\varepsilon(1-\varepsilon)}/\delta/2} = \frac{\delta}{2}.$$