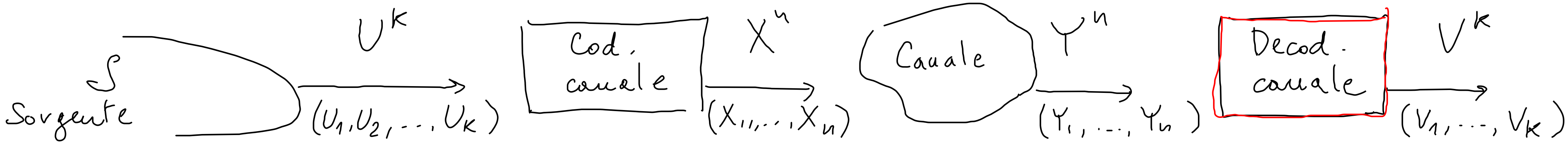


Teorema inverso di Fano (T4.1). Se $R > C/H(S)$, allora la prob. di errore p_{err} rimane limitata inferiormente da una costante positiva p_e^*.

Dimostrazione. $p_{err} = Pr[V^k \neq U^k]$ $\qquad H(S) > C/R$

$R = k/n \qquad C = \max_{p_{X^n}} I(X^n; Y^n)$

Prob. di avere un errore di decodifica in posizione i-esima: $Pr[U_i \neq V_i] =: p_{e/i}$

La prob. $p_{err} = Pr[V^k \neq U^k] \geq \max_{i=1}^{k} Pr[U_i \neq V_i] = \max_{i=1}^{k} p_{e/i} \geq \underbrace{\frac{1}{k} \sum_{i=1}^{k} p_{e/i}}_{p_e \text{ (prob. di errore "media")}}$

Mostreremo che $p_e \geq p_e^* > 0$ (e quindi $p_{err} \geq p_e \geq p_e^* > 0$).

Disuguaglianza di Fano.

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

Dis. di Fano: $\boxed{\text{Se } X \text{ e } \hat{X} \text{ sono v.a., } \Pr[X \neq \hat{X}] =: \varepsilon \text{, e } X \in \mathcal{X} \text{ con } |\mathcal{X}| = K \text{ allora } H(X|\hat{X}) \leq h_2(\varepsilon) + \varepsilon \log(K-1).}$

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

$\Rightarrow$ ottengo $\boxed{H(U_i|V_i) \leq h_2(P_{e/i}) + P_{e/i} \cdot \log(q-1).}$

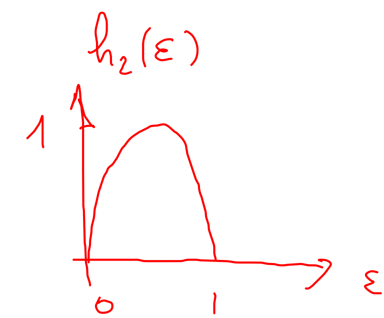
Per le sequenze U^K e V^K, ho per la regola della catena,

$$H(U^K|V^K) = H(U_1|V^K) + H(U_2|U_1, V^K) + H(U_3|U_1, U_2, V^K) + \dots =$$

$$= \sum_{i=1}^{K} H(U_i|U_1, \dots U_{i-1}, V^K) \leq \sum_{i=1}^{K} H(U_i|V_i) \leq$$

$$\leq \sum_{i=1}^{K} \left[h_2(P_{e/i}) + P_{e/i} \cdot \log(q-1) \right]$$

$$\frac{1}{K} H(U^K|V^K) \leq \underbrace{\frac{1}{K} \sum_{i=1}^{K} h_2(P_{e/i})}_{\text{Vorrei esprimere questo termine in funzione di } P_e \text{ anziché delle } P_{e/i}.} + \overset{= P_e}{\boxed{\frac{1}{K} \sum_{i=1}^{K} P_{e/i}}} \log(q-1).$$



Concava

Disuguaglianza di Jensen: Se f è una funz. convessa e Z è una v.a. (reale), allora $E[f(Z)] \geq f(EZ)$.

Se $g = -f$ è concava allora $f = -g$ è convessa. $\longrightarrow$ $\boxed{E[g(Z)] \leq g(EZ)}$

$$\Updownarrow$$

$$E[-g(Z)] \geq -g(EZ)$$

Quindi applichiamo Jensen con $g = h_2(\cdot)$

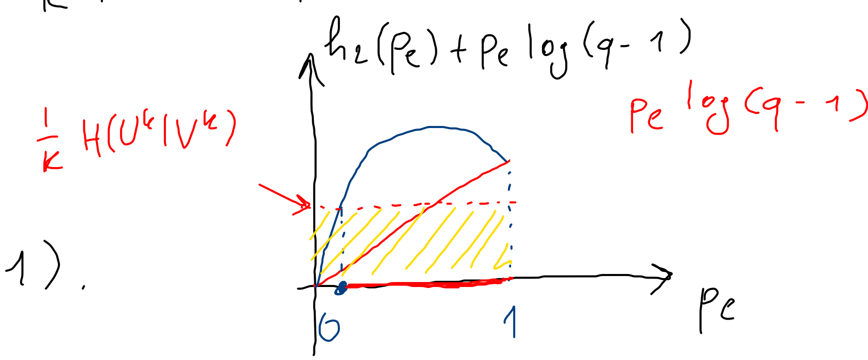
con Z la v.a. pari a $P_{e/i}$ con probabilità $1/K$ per ogni $i = 1, 2, \dots, k$.

Ottengo:

$$E[h_2(Z)] \leq h_2(EZ) \quad . \quad \text{Ma } EZ = \frac{1}{K} P_{e/1} + \frac{1}{K} P_{e/2} + \dots + \frac{1}{K} P_{e/K} =$$

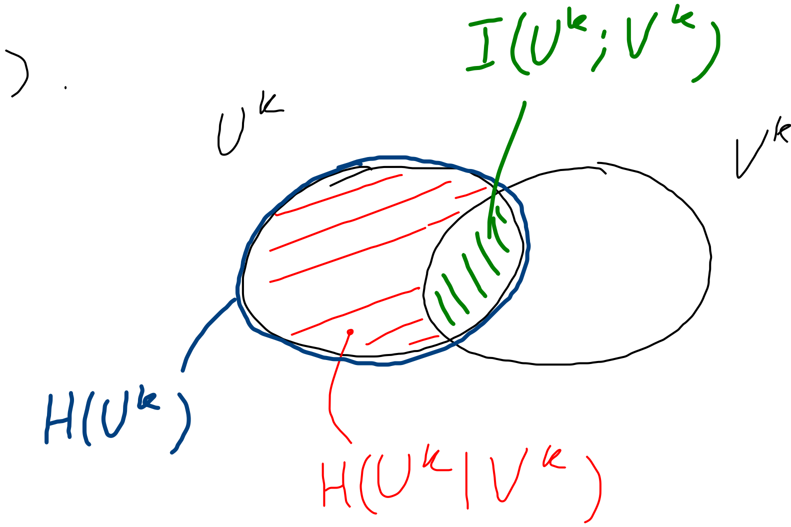
$$= \sum_{i=1}^{K} \frac{1}{K} P_{e/i} = P_e$$

$$\rightarrow \underbrace{E[h_2(Z)]}_{\| \; \frac{1}{K} \sum_{i=1}^{K} h_2(P_{e/i})} \leq h_2(P_e)$$



Ho ottenuto: $\frac{1}{K} H(U^k|V^K) \leq h_2(P_e) + P_e \cdot \log(q-1)$.

$$\frac{1}{k} H(U^k|V^k) \le h_2(p_e) + p_e \cdot \log(q-1).$$



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

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

Le sequenze U^k, X^n, Y^n e V^k sono in catena di Markov: $U^k \to X^n \to Y^n \to V^k$

$$H(U^k) = \sum_{i=1}^{k} H(U_i) = k\,H(\mathcal{S})$$

Per il 2° teorema di elaborazione dati: $I(U^k;V^k) \le I(X^n;Y^n)$

$$\underbrace{U^k \to \overbrace{X^n \to Y^n}^{I(X^n;Y^n)} \to V^k}_{I(U^k;V^k)}$$

Quindi $-I(U^k;V^k) \ge -I(X^n;Y^n)$

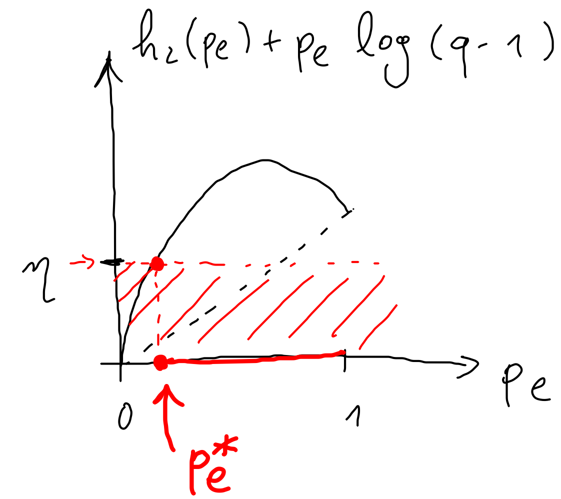
$$\Rightarrow \frac{1}{k} H(U^k|V^k) \ge H(\mathcal{S}) - \frac{1}{k} I(X^n;Y^n)$$

Per il lemma 4.1: $I(X^n;Y^n) \le nC \Rightarrow \frac{1}{k} H(U^k|V^k) \ge H(\mathcal{S}) - \underbrace{\frac{n}{k}}_{=1/R} \cdot C = H(\mathcal{S}) - C/R.$

$$\Rightarrow \quad H(\mathcal{S}) - C/R \le h_2(p_e) + p_e \cdot \log(q-1).$$

Poiché per ipotesi abbiamo $R > C/H(\mathcal{S})$, abbiamo $H(\mathcal{S}) > C/R$

Chiamiamo $\eta = H(\mathcal{S}) - C/R > 0$



Ho $h_2(p_e) + p_e \cdot \log(q-1) \ge \eta$

(qualunque sia il codice di canale)

quindi $p_e \ge p_e^* > 0$

dove p_e^* soddisfa $h_2(p_e^*) + p_e^* \log(q-1) = \eta$.

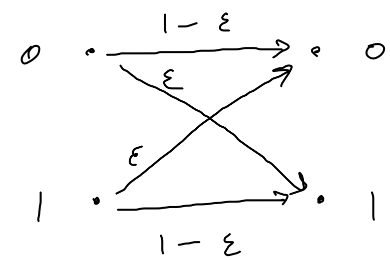
QED.

Caso particolare: se ho una sorgente binaria con $H(\mathcal{S}) = 1$,

il teorema inverso di Fano mi dice che nessun codice di canale con R (tasso) $> C/1$ (capacità) può avere prob. di errore < di p_e^*.

Parte diretta del teorema della codifica di canale

Vedremo la dimostrazione solamente per il canale binario simmetrico $(BSC(\varepsilon))$.



$BSC(\varepsilon)$

Possiamo assumere $\varepsilon < 1/2$

Capacità del canale: $C = 1 - h_2(\varepsilon)$.

- Codice con M di parole di codice $\{x_1, x_2, ..., x_M\}$ (come costruirlo?)
- Decodifica a massima verosimiglianza
- Sorgente (binaria) con d.p. uniforme $(H(S) = 1)$

($\Rightarrow$ le parole di codice $x_1, ..., x_M$ hanno tutte la stessa prob. $1/M$).

$p_e(x_i)$:= prob. di errata decodifica quando la parola di codice inviata è x_i.

$$p_{err} = p_e = \sum_{i=1}^{M} p_e(x_i) \cdot \underbrace{p(x_i)}_{1/M} = \frac{1}{M} \sum_{i=1}^{M} p_e(x_i).$$ (prob. di errore media)

Dimostreremo che per ogni $\delta > 0$ e n sufficientemente grande esiste un codice con tasso $\approx C$ e $p_e < \delta$.
(R)

$$\underbrace{U^k}_{\substack{010..\\ k}} \longrightarrow \underbrace{X^n}_{\substack{001011\\ n}} \longrightarrow \text{Canale} \longrightarrow \underbrace{Y^n}_{\substack{011001\\ n}}$$

$$C/1 = C/H(S)$$

$$\Pr[Y^n = \overbrace{011001}^{y^n} \mid X^n = \overbrace{001011}^{x^n}] = (1-\varepsilon)\varepsilon(1-\varepsilon)(1-\varepsilon)\varepsilon(1-\varepsilon)$$
$$= \varepsilon^2 (1-\varepsilon)^4$$

In generale, $\Pr[Y^n \mid X^n]$ dipende da ε e dalla $\underbrace{\text{distanza di Hamming}}_{\parallel}$ tra y^n e x^n

\# di posizioni in cui i simboli di y^n e x^n differiscono

In generale, se la n-pla di errore ha peso w, la sua probabilità è $\varepsilon^w (1-\varepsilon)^{n-w}$.

x^n	001011
y^n	011001
n-pla di errore	$\boxed{010010}$

← Peso di una sequenza binaria: \# di bit pari a 1

W = numero di bit trasmessi in maniera erronea sul canale

(= peso della n-pla di errore) è una v.a.

con valore atteso $n \cdot \varepsilon$

$$W = \sum_{i=1}^{n} Z_i \qquad \text{dove } Z_i \text{ è } \begin{cases} 1 & \text{se l'i-esimo simbolo è trasmesso erroneamente} \\ 0 & \text{se no.} \end{cases}$$

$$E W = n \underbrace{E Z_i}_{\varepsilon} = n\varepsilon, \qquad Var(W) = Var(Z_1 + \dots + Z_n) \overset{\text{indip.}}{=} \sum_{i=1}^{n} Var(Z_i)$$

$$= \sum_{i=1}^{n} \varepsilon(1-\varepsilon)$$

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