Teoria dell'informazione e della trasmissione

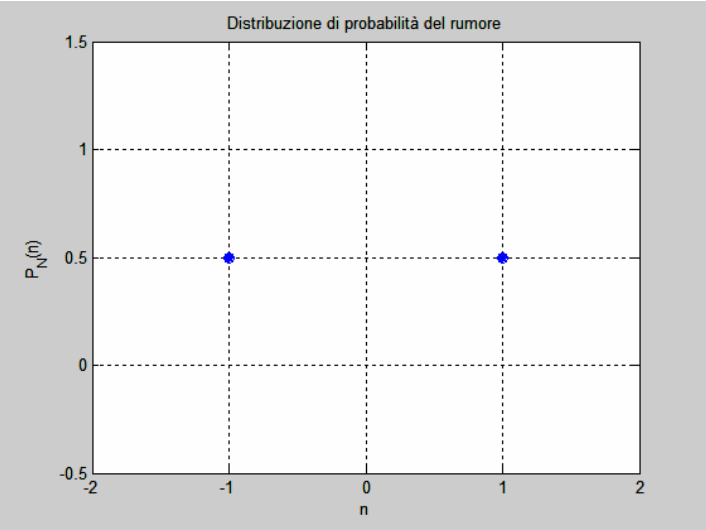
Esercizio: informazione mutua e capacità di canale

Si consideri un canale con ingresso binario $x = \pm 1$ e rumore additivo indipendente (da x) discreto $n = \pm 1$, uniforme ($P_N(n = -1) = P_N(n = +1) = 1/2$).

Calcolare la capacità di canale e spiegare in modo "intuitivo" il valore ottenuto.

Soluzione

La distribuzione di probabilità discreta del rumore N è rappresentata in figura.



Il valore dell'uscita Y, dato che il rumore è additivo, sarà dato dalla somma dei due contributi.

$$Y = X + N$$

Usando il *teorema delle probabilità totali*, possiamo esprimere la probabilità dell'uscita Y nel seguente modo:

$$P_Y(y) = P_{Y|X}(y \mid x = +1)P_X(x = +1) + P_{Y|X}(y)P_X(y \mid x = -1)$$

Ricordiamo il teorema delle probabilità totali: se $A_1, A_2, \ldots, A_n$ sono n eventi tra di loro incompatibili, ma che esauriscono tutti gli eventi possibili (per cui $A_1 \cup A_2 \cup \ldots \cup A_n = S$) e B è un evento qualsiasi, allora:

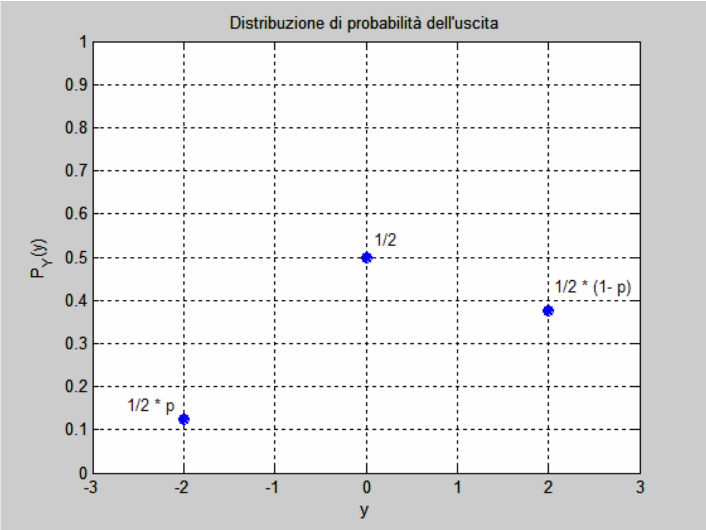
$$P(B) = \sum_{i=1}^{k} P(B / A_i)\,P(A_i)$$

Indicando con p la probabilità $P_X(x = -1)$ e con $(1 - p)$ la probabilità $P_X(x = +1)$, il grafico della distribuzione di probabilità discreta dell'uscita Y avrà la forma in figura.

Come si può osservare, l'uscita può assumere solo tre valori:

- -2 con probabilità $\frac{1}{2}p$ nel caso che l'ingresso valga -1 e il rumore -1;
- 0 con probabilità $\frac{1}{2}$, in quanto può verificarsi con probabilità $\frac{1}{2}p + \frac{1}{2}(1 - p) = \frac{1}{2}$;

- $+2$ con probabilità $\frac{1}{2}(1-p)$, nel caso che l'ingresso valga +1 e il rumore +1.



Il problema chiede di calcolare la capacità di canale (C): questa non è altro che la massima informazione mutua, al variare della distribuzione di probabilità dell'ingresso X.

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

L'informazione mutua può essere definita in due modi:

1. $I(X,Y) = H(X) - H(X \mid Y)$;
2. $I(X,Y) = H(Y) - H(Y \mid X)$.

Per l'esercizio, useremo la seconda definizione.

Ricordiamo la definizione di entropia e di entropia condizionata.

Entropia: $$H(Y) = \sum_i P_Y(y_i) \log \frac{1}{P_Y(y_i)}$$

Entropia condizionata: $$H(Y \mid X) = \sum_j P_X(x_j) \sum_i P_{Y|X}(y_i \mid x_j) \log \frac{1}{P_{Y|X}(y_i \mid x_j)}$$

N.B.: I logaritmi sono in base 2.

Cominciamo a calcolare $H(Y)$.

$$H(Y) = P_Y(y=-2)\log\frac{1}{P_Y(y=-2)} + P_Y(y=0)\log\frac{1}{P_Y(y=0)} + P_Y(y=+2)\log\frac{1}{P_Y(y=+2)}$$

$$H(Y) = \frac{1}{2}p\log\frac{1}{\frac{1}{2}p} + \frac{1}{2} + \frac{1}{2}(1-p)\log\frac{1}{\frac{1}{2}(1-p)}$$

$$H(Y) = \frac{1}{2}\left(p\log\frac{2}{p} + 1 + (1-p)\log\frac{2}{1-p}\right)$$

$$H(Y) = \frac{1}{2}\left(p\left(\log 2 + \log\frac{1}{p}\right) + 1 + (1-p)\left(\log 2 + \log\frac{1}{1-p}\right)\right)$$

$$H(Y)=\frac{1}{2}\left(p\left(1+\log\frac{1}{p}\right)+1+(1-p)\left(1+\log\frac{1}{1-p}\right)\right)$$

$$H(Y)=\frac{1}{2}\left(p+p\log\frac{1}{p}+1+(1-p)+(1-p)\log\frac{1}{1-p}\right)$$

$$H(Y)=\frac{1}{2}\left(\cancel{p}+p\log\frac{1}{p}+2-\cancel{p}+(1-p)\log\frac{1}{1-p}\right)$$

$$H(Y)=\frac{1}{2}\left(2+p\log\frac{1}{p}+(1-p)\log\frac{1}{1-p}\right)$$

$$H(Y)=1+\frac{1}{2}\left(p\log\frac{1}{p}+(1-p)\log\frac{1}{1-p}\right)$$

$$H(Y)=1+\frac{1}{2}H_2[p]$$

Ora calcoliamo $H(Y\mid X)$.

$$H(Y\mid X)=P_X(-1)\left(P_{Y|X}(-2\mid -1)\log\frac{1}{P_{Y|X}(-2\mid -1)}+P_{Y|X}(0\mid -1)\log\frac{1}{P_{Y|X}(0\mid -1)}+P_{Y|X}(+2\mid -1)\log\frac{1}{P_{Y|X}(+2\mid -1)}\right)+$$

$$+P_X(+1)\left(P_{Y|X}(-2\mid +1)\log\frac{1}{P_{Y|X}(-2\mid +1)}+P_{Y|X}(0\mid +1)\log\frac{1}{P_{Y|X}(0\mid +1)}+P_{Y|X}(+2\mid +1)\log\frac{1}{P_{Y|X}(+2\mid +1)}\right)$$

$$H(Y\mid X)=p\left(\frac{1}{2}\log\frac{1}{\frac{1}{2}}+\frac{1}{2}\log\frac{1}{\frac{1}{2}}+0\right)+(1-p)\left(0+\frac{1}{2}\log\frac{1}{\frac{1}{2}}+\frac{1}{2}\log\frac{1}{\frac{1}{2}}\right)$$

$$H(Y\mid X)=p\left(\frac{1}{2}\log 2+\frac{1}{2}\log 2\right)+(1-p)\left(\frac{1}{2}\log 2+\frac{1}{2}\log 2\right)$$

$$H(Y\mid X)=p+1-p=1$$

Il risultato ottenuto è molto interessante: $H(Y\mid X)$ *non dipende* dalla distribuzione di probabilità dell'ingresso. Questo perché il rumore additivo è indipendente dall'ingresso.

L'espressione dell'informazione mutua, da massimizzare, è dunque:

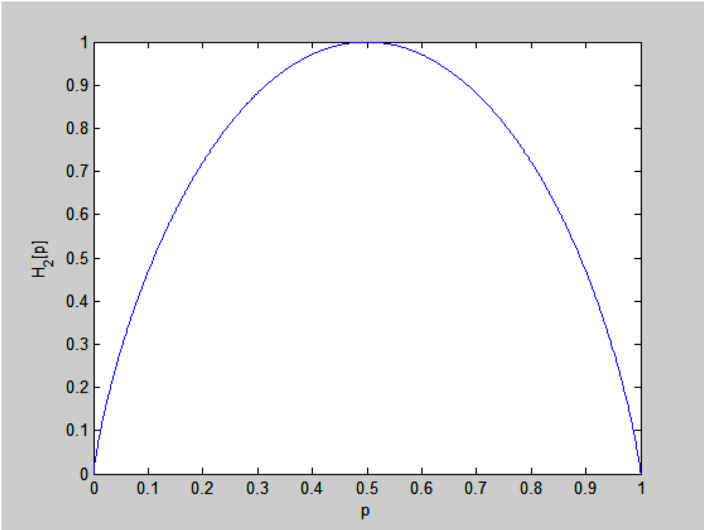
$$I(X,Y)=1+\frac{1}{2}H_2[p]-1=\frac{1}{2}H_2[p]$$

In figura è rappresentato l'andamento di $H_2[p]$ in funzione di *p*. Come già sapevamo, $H_2[p]$ ha un punto di massimo per $p=\frac{1}{2}$ che vale 1.

$$C=\max_{P_X(x)} I(X,Y)=\frac{1}{2}1=\frac{1}{2}\,\text{bit/uso di canale}$$

Abbiamo quindi scoperto che la massima informazione mutua, in questo caso, l'abbiamo quando i due ingressi sono usati con la stessa probabilità.

La capacità di canale è quindi 1/2. La spiegazione intuitiva è che nel 50% dei casi saprò riconoscere esattamente l'ingresso ($y=-2$ e $y=+2$) e nel 50% no ($y=0$).


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
$H_2[p]$
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
p