Mutua informazione

Dati due esperimenti X e Y,

$$I(X;Y) = H(X) - H(X \mid Y) = \sum_{i=1}^{h} \sum_{j=1}^{k} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)}$$

Proprietà:

- $I(X;Y) = I(Y;X)$
- $I(X;Y) = H(X) + H(Y) - H(X,Y)$

È la riduzione di incertezza sul risultato di un esperimento dato dalla conoscenza dell'esito dell'altro, e vale 0 se e solo se X e Y sono indipendenti.

Dimostrazione

$$\begin{aligned}
I(X;Y) &= \sum_{i=1}^{h} \sum_{j=1}^{k} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \\
&= \sum_{i=1}^{h} \sum_{j=1}^{k} p(x_i, y_j) \log \frac{p(y_j)p(x_i \mid y_j)}{p(x_i)p(y_j)} \\
&= \sum_{i=1}^{h} \sum_{j=1}^{k} p(x_i, y_j) \log \frac{p(x_i \mid y_j)}{p(x_i)} \\
&= -\sum_{i=1}^{h} \underbrace{\sum_{j=1}^{k} p(x_i, y_j)}_{p(x_i)} \log p(x_i) + \sum_{j=1}^{k} \sum_{i=1}^{h} \underbrace{p(x_i, y_j)}_{p(y_j)p(y_j \mid x_i)} \log p(x_i \mid y_j) \\
&= H(X) + \sum_{j=1}^{k} p(y_j) \sum_{i=1}^{h} p(x_i \mid y_j) \log p(x_i \mid y_j) \\
&= H(X) - H(X \mid Y).
\end{aligned}$$