Primo teorema di Shannon

Se C è un codice UD su un alfabeto di cardinalità d e una sorgente S, allora

$$L_S(C) \geq \frac{H(S)}{\log d}.$$

Se $d = 2$, $L_S(C) \geq H(S)$.
Questo significa che l'entropia dà un limite inferiore per la compressione: in media non si può codificare un messaggio lungo n con meno di $nH(S)$ bit.

Dimostrazione

Dato che C è UD, allora $K(C) = \sum_{i=1}^{m} d^{-l_i} \leq 1$.
Prendiamo la distribuzione di probabilità:

$$q_i = \frac{d^{-l_i}}{K(C)}$$

ovvero i d^{-l_i} normalizzati. Allora:

$$\begin{aligned}
H(S) &= -\sum_{i=1}^{m} p_i \log p_i \leq -\sum_{i=1}^{m} p_i \log q_i && \text{(lemma del logaritmo)} \\
&= -\sum_{i=1}^{m} p_i \log \frac{d^{-l_i}}{K(C)} = -\sum_{i=1}^{m} p_i \left(\log d^{-l_i} - \log K(C)\right) \\
&= \sum_{i=1}^{m} p_i l_i \log d + \sum_{i=1}^{m} p_i \log K(C) \\
&= L_S(C) \log d + \log K(C) \leq L_S(C) \log d && (K(C) \leq 1)
\end{aligned}$$