

* AVERAGE CODE

LENGTH

would like to characterize
minimum possible average
code length!

$$X \in S = \{s_0, s_1, \dots, s_{n-1}\}$$

$\Pr(s_i) = p_i$

↓
Length of codeword i
 l_i

Average code length

$$= \sum_{i=0}^{M-1} p_i l_i = \bar{L}$$

Kraft Inequality:

$$\sum_{i=0}^{M-1} 2^{-l_i} \leq 1$$

$$q_i = \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}}$$

since $q_i \geq 0$

each $2^{-l_i} \geq 0$

$$\sum_{j=0}^{M-1} 2^{-l_j} \geq 0$$

Further,

$$2^{-l_i} \leq \sum_{j=0}^{M-1} 2^{-l_j}$$

$$\Rightarrow \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}} \leq 1$$

$$\Rightarrow \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}} = q_i \leq 1$$

$$0 \leq q_i \leq 1.$$

Further,

$$\sum_{i=0}^{M-1} q_i$$

$$= \sum_{i=0}^{M-1} \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}}$$

$$= \frac{\sum_{i=0}^{M-1} 2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}} = 1$$

$$0 \leq p_{V_i} \leq 1$$

$$\sum_{i=0}^{M-1} p_{V_i} = 1$$

Probability Distribution.

Kullback-Leibler Divergence

$$p_0, p_1, \dots, p_{n-1}$$
$$q_0, q_1, \dots, q_{n-1}$$

K L Divergence For Probability Mass Functions

$$D(P \parallel Q) \geq 0$$

$$= \sum_{i=0}^{n-1} p_i \log_2 \left(\frac{p_i}{q_i} \right) \geq 0.$$

Replacing
integral by sum

Probability
Densities
by Probability
Masses.

PDFs.

$$D(F_X(x) || g_X(x)) = \int_{-\infty}^{\infty} F_X(x) \log_2 \left(\frac{F_X(x)}{g_X(x)} \right) dx$$

$$D(P || q) \geq 0$$

$$\Rightarrow \sum_{i=0}^{M-1} P_i \log_2 \left(\frac{P_i}{q_i} \right) \geq 0$$

$$\Rightarrow \underbrace{\sum_{i=0}^{M-1} P_i \log_2 P_i}_{-H(X)} + \underbrace{\sum_{i=0}^{M-1} P_i \log_2 \left(\frac{1}{q_i} \right)}_{\geq 0} \geq 0$$

$h_i - e$

$$q_i = \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}}$$

$$\Rightarrow \frac{1}{q_i} = \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}}$$

$$\log_2\left(\frac{1}{q_i}\right) = \log_2 \frac{2^{-l_i}}{\sum_{j=0}^{M-1} 2^{-l_j}}$$

From Kraft Inequality

$$\sum_{j=0}^{M-1} 2^{-l_j} \leq 1$$

$$\Rightarrow \log_2 \sum_{j=0}^{M-1} 2^{-l_j} \leq 0$$

$$= \log_2 2^{l_i} + \log_2 \sum_{j=0}^{M-1} 2^{-l_j}$$

$$= \frac{l_i - \epsilon}{}$$

where $\epsilon \geq 0$.

$$-\epsilon = \log_2 \sum_j 2^{-l_j}$$

From equation (*) above,

$$-H(X) + \sum_{i=0}^{M-1} P_i(l_i - \epsilon) \geq 0$$

$$\sum_{i=0}^{M-1} P_i l_i \geq H(X) + \sum_{i=0}^{M-1} P_i \epsilon$$

$$\underline{\hspace{1cm}} = H(X) + \epsilon \geq 0$$

$$\geq H(X).$$

$$L \geq H(X)$$

$$\sum_{i=0}^{M-1} P_i \log_2 \left(\frac{1}{P_i} \right)$$

Thus we have
a fundamental bound on
average code length
for a given source.

$$L \geq H(X)$$

For any
prefix free
code.

Entropy is a lower bound for average code length.