$$\frac{1}{\log r}\sum_{i=1}^{q} p_i \log\frac{1}{p_i} \leq \sum_{i=1}^{q} p_i l_i \leq \sum_{i=1}^{q} p_i + \frac{1}{\log r}\sum_{i=1}^{q} p_i \log\frac{1}{p_i} \quad \text{——— (1)}$$

We can have $\sum_{i=1}^{q} p_i \log\frac{1}{p_i} = H(S)$;

$$\sum_{i=1}^{q} p_i l_i = L \text{ and}$$

$$\sum_{i=1}^{q} p_i = 1.$$

Then eq (1) becomes:

$$\frac{H(S)}{\log r} \leq L \leq 1 + \frac{H(S)}{\log r}$$

$$\boxed{H_r(S) \leq L \leq 1 + H_r(S)} \quad \text{——— (2)}$$

Eq (2) shows the bounds on optimal code length.

ie $\boxed{L \geq H_r(S)}$ ——— (3)

Eq (2) holds good for nth extension source S to get better efficiency.

ie $S \longrightarrow S^n$ and $L \longrightarrow L_n$.

$$H_r(S^n) \leq L_n \leq 1 + H_r(S^n)$$

ie $$\frac{H(S^n)}{\log r} \leq L_n \leq 1 + \frac{H(S^n)}{\log r} \quad \text{——— (4)}$$

where L_n = average length of codeword of nth extension.

we know that $H(S^n) = n.H(S)$.

Substitute in (4) $\longrightarrow$

$$\frac{n\,H(S)}{\log r} \leq L_n \leq 1 + \frac{n\,H(S)}{\log r}$$

Take $n \cdot H_r(S) \leq L_n \leq 1 + n \cdot H_r(S)$

Divide by n :

$$\boxed{H_r(S) \leq \frac{L_n}{n} \leq 1 + H_r(S)} \quad \text{--- (5)}$$

Taking limit ; $n \longrightarrow \infty$,

$$\lim_{n \to \infty} \frac{L_n}{n} = H_r(S) \leq L .$$

when 'L' is the average length of codewords for the basic source S, then for all values of extension 'n';

$$\frac{L_n}{n} \leq L.$$

Eq (5) is called "Noiseless coding theorem" and it is the basic for Shannon's first theorem or fundamental theorem. Here we are not considering the effect of noise on the code, so termed "noiseless".

SHANNON'S FIRST THEOREM :

Shannon's first theorem or Shannon's fundamental theorem or Noiseless coding theorem states that " given a code alphabet set with 'r' symbols and source alphabet of 'q' symbols, the average lengths of the codewords can be made as close to $H_r(S)$ as possible by increasing the extension "