UNIT-V

SHANNON FANO ENCODING:

Consider a discrete memoryless source with 4 symbols m_1, m_2, m_3, m_4 and the probability of occurrence are 0.25, 0.25, 0.25, 0.25. Find efficiency using Shanon Fano encoding Technique.

m_1	m_2	m_3	m_4
0.25	0.25	0.25	0.25

Write P_k 1 to 4

* probability with decreasing order
* check the condition,

$P(G_1) = P(G_2)$ (or)

$P(G_1) > P(G_2)$

* Algorithm:

1) List the source symbols in decreasing order of their probabilities
2) Divide the given probabilities into 2 groups (G_1, G_2) such that $P[G_1] = P[G_2]$ or $P[G_1] > P[G_2]$
3) Assign the code word [0] for G_1 & [1] for G_2.
4) This process is repeated till further partitioning is not possible.
5) Find $\bar{L}$, $H(x)$, η.

$$H(x) = \sum_{k=1}^{m} P_k \log_2 (1/P_k) \text{. bits/symbol.}$$

$$\bar{L} = \sum_{K=1}^{M} P_k l_k \text{ bits/symbol.}$$

$$\eta = \frac{H(x)}{\bar{L}}$$

where $l_k \rightarrow$ code word length

$M \rightarrow$ number of symbols

P_k	Stage 1	Stage 2	Stage 3	Code word	l_k	$\bar{L} = \sum_{k=1}^{M} P_k l_k$
0.25	0	0	–	00	2	0.5
G_1 0.25	0	1	–	01	2	0.5
0.25	1	0	–	10	2	0.5
G_2 0.25	1	1	–	11	2	0.5
					$\bar{L}$	2 bits/sym

$$H(x) = \sum_{k=1}^{m} P_k \log_2 (1/P_k)$$

$$\log_2 x = \frac{\log_{10} x}{\log_{10} 2} = \frac{\log_{10} x}{0.3010}$$

$$= \frac{1}{0.3010} \left[P_1 \log_{10} (1/P_1) + P_2 \log_{10} (1/P_2) + P_3 \log_{10} (1/P_3) + P_4 \log_{10} (1/P_4) \right]$$

$H(x) =$ 2 bits / symbol.

$$\eta = \frac{H(x)}{\bar{L}} = 2/2 = 100\%$$

(2)

2) Consider a discrete memoryless source emitting 6 symbols $m_1, m_2, m_3, m_4, m_5, m_6$ with their probabilities 0.3, 0.2, 0.25, 0.08, 0.12, 0.05. Find the efficiency, redundancy & code variance using Shanon Fano

P_k	Stage 1	Stage 2	Stage 3	Stage 4	Code word	l_k	$\bar{L}$ (bits/sym)	$H(x) = \sum_{k=1}^{m} P_k \log_2 (1/P_k)$
0.3	0	0	–	–	00	2	0.6	0.5211
0.25	0	1	–	–	01	2	0.5	0.5000
0.2	1	0	–	–	10	2	0.4	0.4644
0.12	1	1	0	–	110	3	0.36	0.3671
0.08	1	1	1	0	1110	4	0.32	0.2915
0.05	1	1	1	1	1111	4	0.2	0.2161
						$\bar{L}$	= 2.38	$H(x)$ = 2.36 bits/symbol

Efficiency $= \frac{H(x)}{\bar{L}} \times 100$

$\eta = \frac{2.36}{2.38} \times 100$

$\eta = 0.991 \times 100$

$\% \eta = 99.1\%$

Redundancy $\gamma = 1 - \eta$

$\gamma = 1 - 0.991$

$\gamma = 0.009$

code Variance :

$\sigma^2 = \sum_{K=1}^{m} P_k (l_k - \bar{L})$

$\sigma^2 = (0.3)(2-2.38) + (0.25)(2-2.38) + (0.2)(2-2.38) + 0.12(3-2.38) + 0.08(4-2.38) + 0.05(4-2.38)$

$\sigma^2 = 0$

(3)

*) HUFFMAN CODING :

•) Algorithm :

1) List out source symbols in decreasing order.
2) The last 2 probabilities are assigned a binary 0 and 1 and the probabilities are added in the first stage.
3) Sum of the probabilities in 1^{st} stage is placed in the II stage such that the probabilities are in the decreasing order.
4) The last 2 probabilities are assigned a binary 0 and 1 and they are added.
5) This procedure is repeated till no further repetition is required.
6) The binary digits are obtained by tracing the path for single message.
7) The code word is obtained by reversing LSB to MSB of traced path.
8) Calculate L, $H(x)$, efficiency

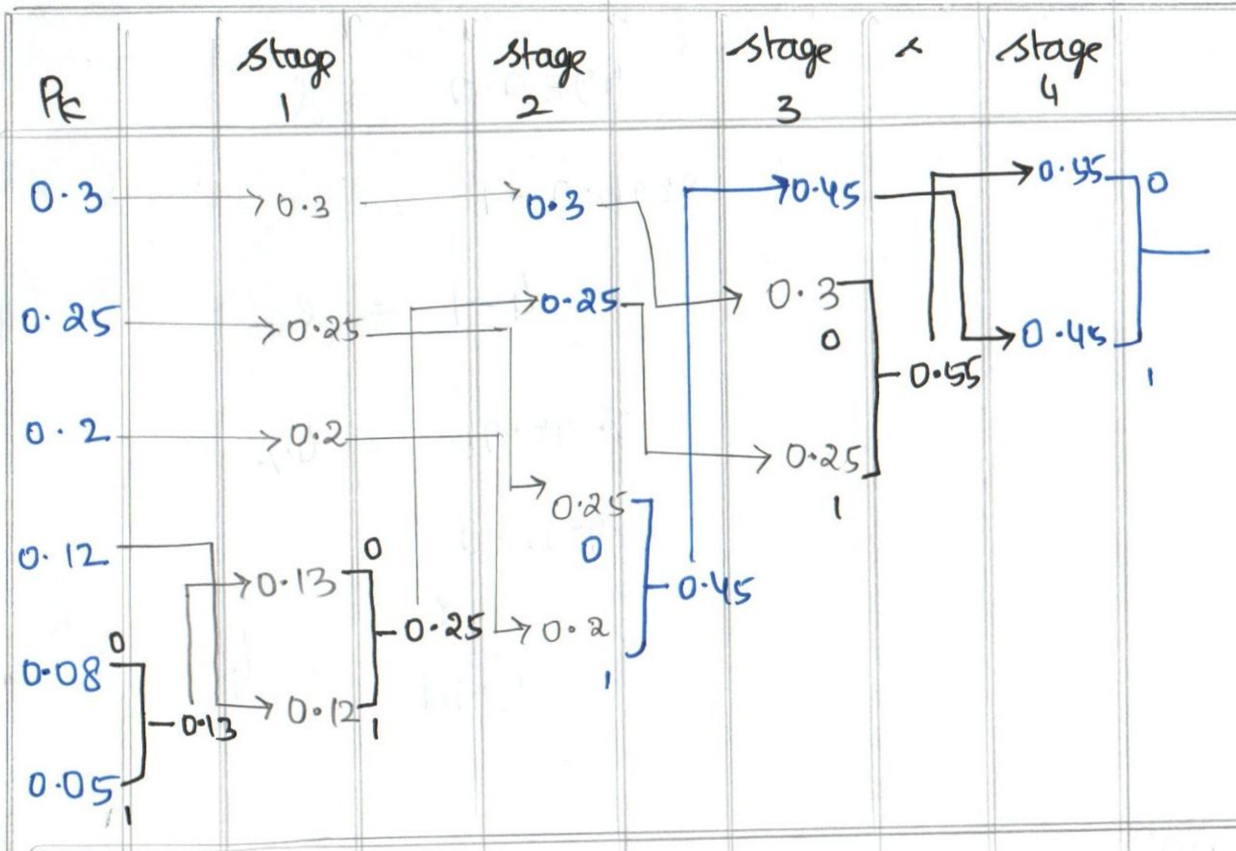
1) Consider a discrete memoryless source emitting 5 symbols.

M_1	M_2	M_3	M_4	M_5
0.4	0.1	0.1	0.2	0.2

. Find ?

2) Consider DMS emitting 6 symbols;

M_1	M_2	M_3	M_4	M_5	M_6
0.05	0.08	0.2	0.12	0.25	0.3



P_k	Rev code word	code word	l_k	$\bar{L}$	$H(x)$
0.3	00	00	2	0.6	0.5211
0.25	01	10	2	0.5	0.5
0.2	11	11	2	0.4	0.4644
0.12	110	011	3	0.36	0.3671
0.08	0010	0100	4	0.32	0.2915
0.05	1010	0101	4	0.2	0.2161

$\underline{\underline{\bar{L} = 2.38}}$ $\boxed{H(x) = 2.36}$ b/s

$$H(x) = \sum_{K=1}^{M} P_K \log_2(1/P_K) \text{ b/s}$$

$$H(x) = P_1 \log_2(1/P_1) + P_2 \log_2(1/P_2) + P_3 \log_2(1/P_3) + P_4 \log_2(1/P_4) + P_5 \log(1/P_5) + P_6 \log_2(1/P_6)$$

$$H(x) = 0.3 \log_2(1/0.3) + 0.25 \log_2(1/0.25) + 0.2 \log_2(1/0.2) + 0.12 \log_2(1/0.12) + 0.08 \log_2(1/0.08) + 0.05 \log_2(1/0.05)$$

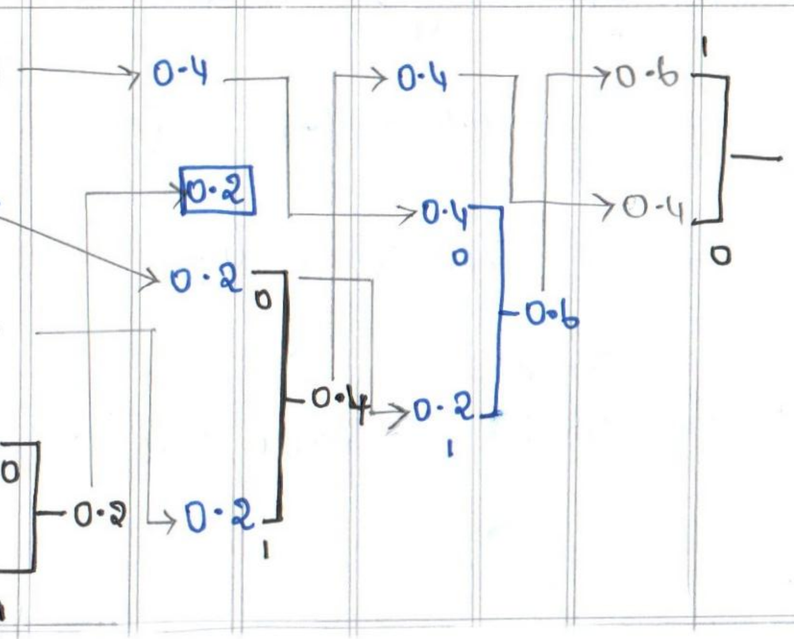
$$\eta = \frac{H(x)}{\bar{L}} = \frac{2.36}{2.38} = 0.991$$

$\% \eta = 99.1\%$

$\gamma = 1 - \eta = 0.008$

$$\sigma^2 = \sum_{K=1}^{M} P_K (l_K - \bar{L}) = 0$$

5



P_k	Reverse Codeword	Code word	l_k	$\bar{\lambda}$	$H(x) = \frac{1}{0.3010}\left[\sum_{k=1}^{m} P_k \log_{10}(1/P_k)\right]$
0.4	00	00	2	0.8	0.528
0.2	01	10	2	0.4	0.464
0.2	11	11	2	0.4	0.464
0.1	010	010	3	0.3	0.332
0.1	110	011	3	0.3	0.332
			$\bar{\lambda}$	2.2	$H(x) = 2.12$ bits/sym

$$\eta = \frac{H(x)}{\bar{\lambda}}$$

$$= 0.9636$$

$$\%\eta = 96.36\,\%$$

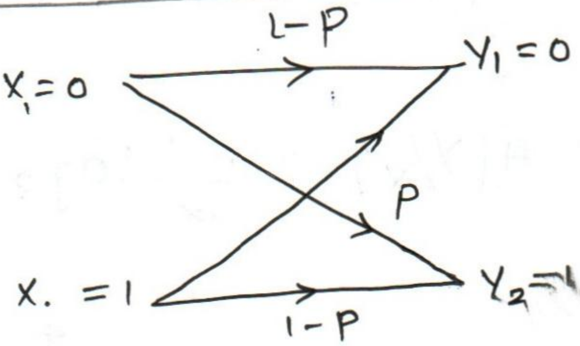
$$\gamma = 1 - \eta$$

$$= 1 - 0.9636$$

$$\gamma = 0.0364$$

6

Channel Capacity for Binary Symmetric Channel.



$$C = \text{Max}\,[\,I(X:Y)\,] \quad \text{— (1)}$$

$$I(X:Y) = H(Y) - H(Y/X) \quad \text{— (2)}$$

Assume $P(X_1) = P(X_2) = 1/2$

$$[P(y_1)\ P(y_2)] = P[(X_1)(X_2)] \begin{bmatrix} P(Y_1/X_1) & P(Y_2/X_1) \\ P(Y_1/X_2) & P(Y_2/X_2) \end{bmatrix}$$

$$= [\,1/2 \quad 1/2\,] \begin{bmatrix} 1-P & P \\ P & 1-P \end{bmatrix}$$

$$= \left[\frac{1-P}{2} + \frac{P}{2} \qquad \frac{P}{2} + \frac{1-P}{2}\right] = [\,1/2 \quad 1/2\,]$$

$$[P(Y_1) \cdot P(Y_2)] = [\,1/2\ \ 1/2\,], \quad \boxed{H(Y) = \sum_{j=1}^{n} P(y_j) \log_2 (1/P.y_j)}$$

by Using H(Y) formula, on substituting

$$= 1/2 \log_2 2 + 1/2 \log_2 2$$

$$\underline{H(Y) = 1}$$

To find $\boxed{H(Y/X) = \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i, y_j) \log_2 \frac{1}{P(y_j/x_i)}}$

$$P(X_1, y_1) = P(y_1/x_1)\, P(X_1) = (1-P)\, 1/2 = (1-P)/2$$

$$P(X_1, Y_2) = P \cdot (1/2) = P/2$$

$$P(X_2, Y_1) = P \cdot (1/2) = P/2$$

$$P(x_2, y_2) = (1-P)\tfrac{1}{2} = 1-P/2 .$$

$$\therefore H(Y/X) = \frac{1-P}{2}\log_2 \frac{1}{(1-P)} + \frac{P}{2}\log_2 1/P + \frac{1-P}{2}\log_2 \frac{1}{(1-P)}$$

$$+ P\log_2 1/P .$$

$$H(Y/X) = \underbrace{1 - P\log_2 1/_{1-P} + P\log_2 1/P}_{H(P)}$$

Consider $H(Y/X) = H(P)$

$$H(Y/X) = H(P)$$

WKT, $C = \max I(X:Y)$

$$I(X:Y) = H(Y) - H(Y/X)$$

$$\therefore I(X:Y) = 1 - H(P) .$$

Definition: Binary Symmetric channel [BSC]

The binary Communication channel is Symmetric. When the probability of receiving '1' if a is '0' is sent is the same probability of receiving a '0' if a '1' is Sent. $P(y_1/x_2) = P(y_2/x_2)$

Information capacity theorem

(or) Shannon-Hartley theorem (or) continuous capacity theorem

The information capacity of continuous channel of B Hz disturbed by AWGN of zero mean and PSD $N_o/2$ and limited to a B.W of B Hz is given as

$$C = B \log_2 \left(1 + \frac{P}{N_o B}\right) \text{ bits/second}$$

$$C = B \log_2 \left(1 + \frac{P}{N_o B}\right) = B \log_2 \left(1 + \frac{S}{N}\right) \text{ bits/second}$$

$$C = \max \; I(x:y)$$

$$C = [H(x) - H(x/y)]$$

channel $\eta = \frac{\text{actual transmission}}{\text{max transmission}} = \frac{I(x:y)}{\max[I(x:y)]}$

$$\boxed{\eta = I(x:y)/C}$$

$$H(x) = B \log_2 [2\pi e (s + \sigma^2)] \quad - ③$$

$$H(x/y) = B \log_2 [2\pi e \sigma^2] \quad - ④$$

Substitute ③ & ④ in ①,

$$C = B \log_2 [2\pi e [s + \sigma^2]] - B \log_2 [2\pi e \sigma^2]$$

$$= B \log_2 [2\pi e [S+N]] - B \log_2 [2\pi e N]$$

$$C = B \log_2 \frac{2\pi e (S+N)}{2\pi e N} \text{ bits/sec}$$

$$C = B \log_2 \left[\frac{S+N}{N}\right]$$

$$C = B \log_2 \left[1 + \frac{S}{N}\right] \text{ bits/sec}$$

Trade OFF between bandwidth and signal to noise ratio:

$$C = B\log_2\left[1+\frac{S}{N}\right] - ①$$

Case (i): $N=0$, $C\to\infty$

$$C = B\log_2\left[1+\frac{S}{0}\right] \Rightarrow C \doteq B\log_2[1+\infty] \Rightarrow \boxed{C=\infty}$$

Case (ii): If $B\to\infty$, $C = 1.44\frac{S}{N}$

upper limit calculation:-

$$N = N_0 B \quad -②$$

Subs ② in ①,

$$C = B\log_2\left(1+\frac{S}{N_0 B}\right) - ③$$

Multiply & divide by $\frac{S}{N_0}$ ③ eqn,

$$C = B\frac{S}{N_0}\frac{N_0}{S}\log_2\left(1+\frac{S}{N_0 B}\right) - ④$$

$$= \frac{S}{N_0}\log_2\left(1+\frac{S}{N_0 B}\right)\frac{N_0 B}{S}$$

$$= \frac{S}{N_0}\log_2(1+x)\frac{1}{x}$$

$x = \frac{S}{N_0 B}$; $N_0 B\to\infty$, $x=0$.

$$C = \frac{S}{N_0 B}\lim_{S\to 0}\log(1+x)^{1/x}$$

$$\lim_{S\to 0}(1+x)^{1/x} = e \Rightarrow C = \frac{S}{N_0}\log_2 e \Rightarrow C = \frac{S}{N_0}\frac{\log_{10} e}{\log_{10} 2}$$

$$\Rightarrow \boxed{C = 1.44\frac{S}{N_0}}$$

Properties of Entropy

(i) $H=0$ if $P_K=0$ (or) $P_K=1$

Entropy is Zero for Surely Occurred event and impossible event.

proof :-

$H=0$ if $P_K=0$ (or) $P_K=1$

when $P_K=0$ (ie) event is impossible

W.K.T $H = \sum_{K=1}^{K} P_K \log_2 (1/P_K)$

$= \sum_{K=0}^{K} 0 \log_2 (1/0)$

$\boxed{H=0}$

when $P_K=1$, (ie) event is sure

$H = \sum_{K=1}^{K} (1) \log (1/1)$

$\boxed{H=0}$

Example For Property 2:

Eg: Let $P_1=0.5$, $P_2=0.5$.

$H(x) = \sum_{K=1}^{M} P_K \log_2 (1/P_K)$

$= \frac{1}{0.301} \sum_{K=1}^{2} P_K \log_2 (1/P_K)$

$= \frac{2}{0.301} \left[0.5 \log\left(\frac{1}{0.5}\right)\right]$

$H(x) = 1.000$ b/s $\longrightarrow$ 1

$H(x) = \log_2 m$

$H(x) = \log_2 2$

$H(x) = 1.000$ bits/sec $\longrightarrow$ 2

Since ① = ②

$H(x) = \log_2 m$ (when all the symbols have same probability).

(ii) $H = \log_2 K$, when the symbols are equally likely for K symbols, i.e $P_K = 1/K$

Let us consider symbols are 0.25, 0.25, 0.25, 0.25

$K \rightarrow$ No of symbols are equally likely

Hence where $K = 4$ & $0.25 = 1/4$

$$P_K = 1/K$$

$$H = \sum_{K=1}^{K} P_K \log_2 (1/P_K)$$

$$= \frac{1}{4}\log_2\left(\frac{1}{1/4}\right) + \frac{1}{4}\log_2\left(\frac{1}{1/4}\right) + \frac{1}{4}\log_2\left(\frac{1}{1/4}\right)$$

$$+ \frac{1}{4}\log_2(1/4)$$

$$= \frac{4}{4}\log_2 4$$

$$\boxed{H = \log_2 4}$$

i.e, $H = \log_2 K$ when K is the no of symbols are equally likely.

(iii) Upper bound on entropy is given as $H_{max} \leq \log_2 M$

we know that, $\log_2 M \geq H(x) \geq 0$

Consider any two probability distributions $(P_1, P_2, P_3, \ldots P_M)$ & $(q_1, q_2 \ldots q_m)$ on the alphabet $X = \{x_1, x_2, \ldots x_M\}$ of a DMS source.

Then $\sum_{K=1}^{M} P_K \log_2\left(\frac{q_K}{P_K}\right) \Rightarrow \sum_{K=1}^{M} \frac{P_K \log_{10}\left(\frac{q_K}{P_K}\right)}{\log_{10} 2}$

property of natural log.

$$\log x \leq x-1 \,,\; x>0$$

$$\therefore \sum_{K=1}^{M} \frac{P_K \log_{10}\left(\frac{q_K}{P_K}\right)}{\log_{10} 2} \leq \sum_{K=1}^{M} \frac{1}{\log_{10} 2} P_K\left(\frac{q_K}{P_K}-1\right)$$

$$\leq \frac{1}{\log_{10} 2} \sum_{K=1}^{M} (q_K - P_K)$$

$$\leq \frac{1}{\log_{10} 2} \left(\sum_{K=1}^{M} q_K - \sum_{K=1}^{M} P_K\right)$$

W.K.T $\sum_{K=1}^{M} P_K = \sum_{K=1}^{M} q_K = 1$

$$\boxed{\therefore \sum_{K=1}^{M} P_K \log_2 \frac{q_K}{P_K} \leq 0}$$

$$\log ab = \log a + \log b$$

$$\therefore \sum_{K=1}^{M} P_K \log_2 q_K + \sum_{K=1}^{M} P_K \log_2 \frac{1}{P_K} \leq 0$$

$$\sum_{K=1}^{M} P_K \log_2 \frac{1}{P_K} \leq \sum_{K=1}^{M} P_K \log_2 \frac{1}{q_K}$$

Sub $q_K = \frac{1}{M}$

$$\leq \sum_{K=1}^{M} P_K \log_2\left(\frac{1}{q_K}\right)$$

$$\leq \sum_{K=1}^{M} P_K \log_2 M$$

$$\leq \log_2 M \sum_{K=1}^{M} P_K$$

$$\sum_{K=1}^{M} P_K \log_2 \frac{1}{P_K} \leq \log_2 M$$

$$\boxed{H \leq \log_2 M}$$