Formulae :-

$$I(E_K) = \log_2 \frac{1}{P_K}$$

$$I(E_K) = -\log P_K$$

Entropy :-

$$H(x) = -\sum_{K=0}^{K-1} P_K \log P_K$$

Q- What is the entropy of a source having a symbol P_K probability $P_K = 0.5, 0.3, 0.15, 0.05$.

Sol- $$H(x) = -\sum_{K=0}^{K-1} P_K \log P_K$$

$$= -\sum_{K=0}^{4-1} P_K \log P_K = -\sum_{K=0}^{3} P_K \log P_K$$

$$= -[0.5 \log_2 0.5 + 0.3 \log_2 0.3 + 0.15 \log_2 0.15 + 0.05 \log_2 0.05]$$

$$\boxed{\text{Note} - \log_2 3 = \frac{\cancel{\log} 3}{\cancel{\log} 2} \quad \frac{\ln 3}{\ln 2}}$$

$$\therefore E(X) = -[0.5 \times -1 + 0.3(-1.73) + 0.15(-2.73) + 0.05(-4.3)]$$

$$= -[-0.50 - 0.519 - 0.40 - 0.215]$$

$$= 1.634 \text{ bits}$$

Q Find entropy

a	b	c	d
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

Sol-

$$a = -\log \frac{1}{2} = -\frac{\ln 1/2}{\ln 2} = 1 \times \frac{1}{2} = \frac{1}{2}$$

$$b = -\frac{1}{4} \log \frac{1}{4} = \frac{1}{2}$$

$$c = -\frac{1}{8} \log \frac{1}{8} = 0.375 = d$$

$$H(X) = \frac{1}{2} + \frac{1}{2} + 0.375 + 0.375$$

$$= \frac{7}{4} \text{ bits}$$

Q- Three probability P_1, P_2 and P_3

$$H(x) = -P_1 \log_2 P_1 - P_2 \log_2 P_2 - P_3 \log_2 P_3$$

As we Know

$$P_1 + P_2 + P_3 = 1$$

$$\Rightarrow P_3 = 1 - P_1 - P_2$$

$$\therefore H(x) = -P_1 \log_2 P_1 - P_2 \log_2 P_2 - (1 - P_1 - P_2) \log(1 - P_1 - P_2)$$

→ Properties of entropy

i) $H(x)$ is non-negative i.e. $H(x) \geq 0$

ii) $H(x) = 0$; $P = 0$

iii) when probability is equally distributed then the entropy $H(x)$ is maximum.

Put

$$P_1 = K_1$$

$$P_2 = \frac{1 - K_1}{2}$$

$$P_2 = P_3$$

$$\therefore H(x) = -K_1 \log K_1 - \left(\frac{1-K_1}{2}\right) \log\left(\frac{1-K_1}{2}\right) - \left(1 - K_1 - \left[\frac{1-K_1}{2}\right]\right) \log\left[1 - K_1 - \left(\frac{1-K_1}{2}\right)\right]$$

$$H(x) = -\left\{K_1 \log K_1 + \left(\frac{1-K_1}{2}\right)\log\left(\frac{1-K_1}{2}\right)\right.$$

$$\left. + \left(\frac{1-K_1}{2}\right)\log\left(\frac{1-K_1}{2}\right)\right\}$$

$$H(x) = -\left\{K_1 \log K_1 + \frac{2}{2}(1-K_1)\log\left(\frac{1-K_1}{2}\right)\right\}$$

$$= -\left\{K_1 \log K_1 + (1-K_1)\log\frac{(1-K_1)}{2}\right\}$$

diff w. r. to K_1

$$\frac{dH(x)}{dK_1} = -\left\{K_1 \times \frac{1}{K_1 \ln 2} + \log K_1 + (-1)\log\left(\frac{1-K_1}{2}\right)\right.$$

$$\left. + (1-K_1)\frac{\cancel{2}}{(1-K_1)\ln 2} \cdot \frac{1}{\cancel{2}}(-1)\right\}$$

$$= -\left\{\cancel{\frac{1}{\ln 2}} + \log K_1 - \log\left(\frac{1-K_1}{2}\right)\right.$$

$$\left. - \cancel{\frac{1}{\ln 2}}\right\}$$

$$\boxed{\begin{array}{l} \cancel{= -\left\{-\frac{1}{\ln 2} + \log K_1 - \log\left(\frac{1-K_1}{2}\right)\right\}} \\ \cancel{= \frac{1}{\ln 2} \log K_1 + \log\left(\frac{1-K_1}{2}\right)} \end{array}}$$

For max entropy $\frac{dH(x)}{dK_1} = 0$

$$\therefore \log_2 K_1 - \log\left(\frac{1-K_1}{2}\right) = 0$$

$$\log K_1 = \log\left(\frac{1-K_1}{2}\right)$$

$$K_1 = \frac{1-K_1}{2}$$

$$2K_1 = 1 - K_1$$

$$2K_1 + K_1 = 1$$

$$3K_1 = 1$$

$$\Rightarrow K_1 = \frac{1}{3}$$

$$\therefore P_1 = \frac{1}{3}$$

$$P_2 = \frac{1-K_1}{2} = \frac{1-\frac{1}{3}}{2}$$

$$= \frac{\frac{3-1}{3}}{2} = \frac{\frac{2}{3}}{2}$$

$$= \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$$

$$P_3 = P_2 = \frac{1}{3}$$

Note:-

$$\frac{d}{dx}\log_2 x = \frac{1}{x \ln 2}$$

$$\frac{d}{dx}\log_b x = \frac{1}{x \ln b}$$

→ Properties of entropy

iv) $H(X) \leq \log_b M$ [M = no. of bits]

v) $H(X) = \log_b M$

If and only if all the probabilities are equal so that $P(x_i) = P_i = \frac{1}{M}$ for all i.

Let us consider the probabilities P_i and q_i where $i = 1, 2, 3, \ldots m$.

$$\sum_{i=1}^{M} p_i = \sum_{i=1}^{M} q_i = 1$$

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

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

$$\sum_{i=1}^{M} P_i \log \frac{P_i}{1/M} \geq \ln b \sum_{i=1}^{M} P_i \left(1 - \frac{1}{MP_i}\right) = 0$$

$$\sum_{i=1}^{M} P_i \log MP_i \geq \ln b \sum_{i=1}^{M} P_i \left(1 - \frac{1}{MP_i}\right) = 0$$

$$\geq \ln b \sum_{i=1}^{M} P_i - \ln b \sum_{i=1}^{M} \cancel{P_i} \frac{1}{M\cancel{P_i}}$$

$$\geq \ln b \sum_{i=1}^{M} P_i - \ln b \sum_{i=1}^{M} \frac{1}{M}$$

when $P_i = q_i$

$$\sum_{i=1}^{M} P_i \log MP_i \geq 0$$

$$\sum_{i=1}^{M} P_i \log_b P_i + \sum_{i=1}^{M} \log_b MP_i \geq 0$$

$$\sum_{i=1}^{M} P_i \log_b P_i + M \sum_{i=1}^{M} \log_b P_i \geq 0$$

$$\sum_{i=1}^{M} P_i \log_b P_i \geq - \sum_{i=1}^{M} M \log P_i$$

$$\sum_{i=1}^{M} P_i \log_b P_i \leq \log_b M$$

$$\boxed{H(X) \leq \log_b M}$$

Note :-

$$\log_b 2 = \frac{\ln 2}{\ln b}$$

Q- ~~$H(x) = -P_i \Sigma \log$~~

$$H(X) = -\sum_{K=0}^{5} P_i \log_2 P_i$$

$$= -\sum_{K=0}^{5} \frac{1}{6} \log_2 \frac{1}{6}$$

$$\cancel{6} \times \frac{1}{\cancel{6}} \log_2 6$$

$$= \log_2 6$$

$$= 2.58$$

Q- There are 5 boys and 4 girls sit in a bench. Find what are the possible way to sit under a condition 4 girls sit together.

Sol- $6! \times 4! = 17280$

* Joint And Conditional Entropy

Take x_i as source alphabate with M sampl
y_i as destination alphabate with
N samples.

$$H(x) = -\sum_{i=1}^{M} P(x_i) \log_2 P(x_i)$$

$$H(y) = -\sum_{j=1}^{N} P(y_i) \log_2 P(y_i)$$

∴ Joined entropy

$$H(x,y) = -\sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log_2 P(x_i, y_j)$$

If x_i and y_i are statistically
independent then

$$P(x_i, y_i) = P(x_i) P(y_i)$$

$$\sum_{j=1}^{N} P(x_i, y_j) = P(x_i)$$

$$\sum_{i=1}^{M} P(x_i, y_i) = P(y_j)$$

$$H(X,Y) = -\sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j) \log_2 P(x_i) P(y_i)$$

$$= -\sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_i) \log_2 P(x_i)$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_i) \log_2 P(y_i)$$

If X and Y are not independent then

$$H(X,Y) = -\sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_i) \log_2 \left[P(y_i/x_i) P(x_i)\right]$$

$$= -\sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j) \log P(x_i)$$

$$- \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j) \log_2 P(y_i/x_i)$$

$$H(X,Y) = H(X) - \sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j) \log P\left(\frac{y_i}{x_i}\right)$$

Last term is called as conditional entropy.

$$H(X,Y) = H(X) + H\left(\frac{Y}{X}\right)$$

or,

$$H(X,Y) = H(Y) + H(X/Y)$$

$$0 \leq H\left(\frac{Y}{X}\right) \leq H(Y)$$

$$0 \leq H\left(\frac{X}{Y}\right) \leq H(X)$$

★ Mutual Information

It is defined the amount of information transfered when x_i is transmitted & y_i is received.

$$I(x_i, y_j) = \log_2 \frac{P(x_i/y_j)}{P(x_i)} \text{ bits}$$

or

$$= \log_2 \frac{P(x_i, y_j)}{P(x_i)P(y_j)}$$

★ Avg. mutual information

Represented by

$$I(X,Y) = \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log_2 \frac{P(x_i/y_j)}{P(x_i)}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i, y_j) \log_2 \frac{P(x_i, y_j)}{P(x_i)P(y_j)}$$

$$I(X,Y) = \sum_{i=1}^{M} \sum_{j=1}^{N} \{P(x_i, y_j)[\log_2 P(x_i y_j) - \log_2 P(x_i)\, P(y_j)]\}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i y_j) \log_2 P(x_i y_j) - \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i y_j) \log_2 P(x_i) - \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i y_j) \log_2 P(y_j)$$

$$I(X,Y) = -H(X,Y) + H(X) + H(Y) \geq 0$$

$$= -\cancel{H(X)} - H(Y/X) + \cancel{H(X)} + H(Y) \geq 0$$

$$I(X,Y) = H(Y) - H(Y/X) \geq 0$$

||ly

$$I(X,Y) = H(X) - H(X/Y) \geq 0$$

$$\boxed{H(X,Y) \leq H(X) + H(Y)}$$

★ Properties of mutual information

i) $I(X,Y) = I(Y,X)$

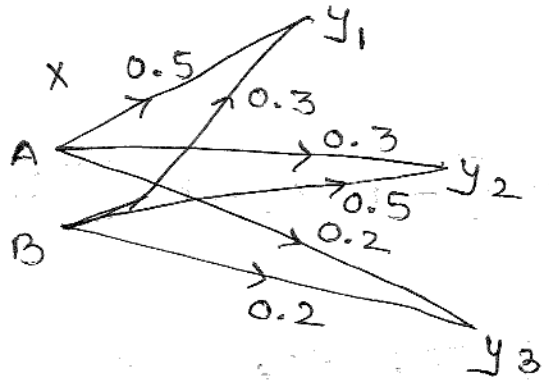
<u>Proof</u>

$$I(X,Y) = \sum_{i=1}^{M} \sum_{j=1}^{N} P(x_i y_j) \log_2 \frac{P(x_i/y_j)}{P(x_i)}$$

$$P(x_i, y_j) = P\left(\frac{x_i}{y_j}\right) P(y_j) = P\left(\frac{y_j}{x_i}\right) P(x_i)$$

$$\log_2 \frac{P(x_i/y_j)}{P(x_i)} = \log_2 \frac{P(y_j/x_i)}{P(y_j)}$$

Q- Find the entropy



Sol- $P(Y/X) =$

	Y_1	Y_2	Y_3
A	0.5	0.3	0.2
B	0.3	0.5	0.2
	0.8	0.8	0.4

~~$P(Y_1) = 0.8$ $P(Y_2) = 0.8$ $P(Y_3) = 0.4$~~

~~$P(A) = \frac{1}{2}$ $P(B) = \frac{1}{2}$~~

Sol- $P(Y/X) =$

	Y_1	Y_2	Y_3
A	0.5	0.3	0.2
B	0.3	0.5	0.2
	0.8	0.8	0.4

$P(Y) = 2$

$P(A) = \frac{1}{2}$ $P(B) = \frac{1}{2}$

$\therefore P(X, Y) =$

	Y_1	Y_2	Y_3
0.5	0.25	0.15	0.10
0.5	0.15	0.25	0.10
	0.40	0.40	0.20

$$P(Y) = P(Y_1) + P(Y_2) + P(Y_3)$$
$$= 0.4 + 0.4 + 0.2$$
$$= 1$$

$$H(Y) = -\sum_{i=1}^{3} P(y_i) \log_2 P(y_i)$$

$$= -P(Y_1)\log_2 P(Y_1) - P(Y_2)\log_2 P(Y_2) - P(Y_3)\log_2 P(Y_3)$$

$$= -0.4\,\frac{\ln 0.4}{\ln 2} - 0.4\,\frac{\ln 0.4}{\ln 2}$$

$$-0.2\,\frac{\ln 0.2}{\ln 2}$$

$$= 1.52193$$

$$H(Y/X) = -\sum_{i=1}^{M}\sum_{j=1}^{N} P(x_i, y_j)\log P(y/x_i)$$

$$= -0.25 \times \log_2 0.5 - 0.15 \log_2 0.3$$

$$-0.10 \log_2 0.2 - 0.15 \log_2 0.3$$

$$-0.25 \log_2 0.5 - 0.10 \log_2 0.2$$

$$= 1.485$$

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

$$= 1.52 - 1.485$$

$$= 0.036$$

<Ans>

i)

0.3

3

0.2

→ One dimentional entropy in continuous case :-

- $$H(X) = -\int_{-\infty}^{\infty} f(x)\log_2 f(x)\,dx$$

- $$H(X) = E[-\log f(x,y)]$$
$$= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\log_2 f(x,y)\,dx\,dy$$

- $$H(Y) = -\int_{-\infty}^{\infty} f(y)\log_2 f(y)\,dy$$

- $$H(Y/X) = -\iint f(x,y)\log_2 \frac{f(x,y)}{f(x)}\,dx\,dy$$

- $$H(X/Y) = -\iint f(x,y)\log_2 \frac{f(x,y)}{f(y)}\,dy\,dx$$

- $$H(X_1, X_2, X_3, X_4 \cdots X_n) = -\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1, x_2 \cdots x_n)\log f(x_1, x_2, x_3 \cdots x_n)\,dx_1,\,dx_2 \cdots\cdots dn$$

- $P(x_i) \approx W_x(x_i)\,\Delta x_i$

$$H(x) \overset{\lim_{\Delta x \to 0}}{=} -\sum P(x_i)\log_2 P(x_i)$$

$$\overset{\lim_{\Delta x \to 0}}{=} -\sum W_x(x_i)\,\Delta x_i \log_2 W_x(x_i)\,\Delta x_i$$

$$H(x) = -\mathop{\mathrm{lt}}_{\Delta x \to 0} \sum W_x(x_i)\,\Delta x_i \log_2 W_x(x_i)$$

$$- \mathop{\mathrm{lt}}_{\Delta x \to 0} \sum W_x(x_i)\,\Delta x_i \log_2 \Delta x_i$$

$$= -\int_{-\infty}^{\infty} W_x(x_i)\log_2 W_x(x_i)\,dx - \underbrace{\mathop{\mathrm{lt}}_{\Delta x \to 0} \log_2 \Delta x_i}_{\infty}$$

$$\underbrace{\int_{-\infty}^{\infty} W_x(x_i)\,dx}_{1}$$

$$H(x) = -\int_{-\infty}^{\infty} W_x(x_i)\log_2 W_x(x_i)\,dx$$

Q ~~Define six~~ Given six messanger m_1, m_2, m_3 m_4, m_5, m_6. The probabilities are

	m_1	m_2	m_3	m_4	m_5	m_6
Δx_i	$\frac{1}{3}$	$1/4$	$1/8$	$1/8$	$1/12$	$1/12$

Find the shannon fannon. Find coding efficiency and retundancy.

			code word	length
Sol-	m_1	$1/3 = 0.37$	0 0	2
	m_2	$1/4 = 0.25$	0 1	2
	m_3	$1/8 = 0.125$	1 0 0	3
	m_4	$1/8 = 0.125$	1 0 1	3
	m_5	$1/12 = 0.08$	1 1 0	3
	m_6	$1/12 = 0.08$	1 1 1	3

Entropy $H(X) = -\cancel{P_i \sum_{i=1}} \sum_{i=0}^{5} P(X_i) \log_2 P(X_i)$

$= -m_1 \log_2 m_1 - m_2 \log_2 m_2 - m_3 \log_2 m_3$

$- m_4 \log_2 m_4 - m_5 \log_2 m_5 - m_6 \log_2 m_6$

$$H(x) = 2.356 = 2.378$$

$$\bar{L} = P_1L_1 + P_2L_2 + P_3L_3 + P_4L_4 + P_5L_5 + P_6L_6$$

$$= 0.66 + 0.5 + 0.375 + 0.375 + 0.25 + 0.25$$

$$= 2.416$$

$$\eta = \frac{H(x)}{\bar{L}\log D} = \frac{2.3758}{2.416 \log_2 2}$$

$$= 0.985$$

$$R_c = 1 - \eta = 1 - 0.985 = 0.015$$

Q- Given messages are 9. The probabilities are

			codeword	Length
m_1	0.49	1	0	1
m_2	0.14	3	1 0 0	3
m_3	0.14	2	1 0 1	3
m_4	0.07	5	1 1 0 0	4
m_5	0.07	4	1 1 0 1	4

m_6	0.04	1110	4
	------ 6		
m_7	0.02	11110	5
	------ 7		
m_8	0.02	111110	6
	------ 8		
m_9	0.01	111111	6

$$H(X) = -\sum_{i=0}^{8} P(x_i) \log_2 P(x_i)$$

$$= -0.49 \frac{\ln 0.49}{\ln 2} - 0.14 \frac{\ln 0.14}{\ln 2} - 0.14 \frac{\ln 0.14}{\ln 2}$$

$$- 0.07 \times 2 \frac{\ln 0.07}{\ln 2} - 0.04 \frac{\ln 0.04}{\ln 2}$$

$$- 0.02 \times 2 \frac{\ln 0.02}{\ln 2} - 0.01 \frac{\ln 0.01}{\ln 2}$$

$$= 2.313$$

$$\bar{L} = 0.49 + 0.42 + 0.42 + 0.28 + 0.28 + 0.16 + 0.12 + 0.1 + 0.06$$

$$= 2.33$$

$$\eta = \frac{2.313}{2.33 \log_2 2} = 0.992$$

$$R_c = 1 - 0.992 = 0.01$$