: Information theory :

Information souce may be an object which produces an event, the outcome of which is selected at random according to a probability distrubution. The analog souce can be transformed to discrete souce through the use of sampling & quantization techniques. The set of souce involves is called the souce alphabet and the elements of the set are called symbols or letters.

Information content of a discrete memory less souce propeties

1) Information should be proportional to uncertainity of outcome
2) Information containing independents outcomes should add to get overall information code.

Information content of symbol [logarithmic measure of information] let us consider a Discrete memory less souce (DMS) $X = \{x_1, x_2, x_3 \cdots x_m\}$. The information content of a symbol x_i is denoted by $I(x_i)$ is defined by $I(x_i) = \log_b \frac{1}{P(x_i)}$

$P(x_i) \rightarrow$ probability of occurence of $x(i)$

$\Rightarrow I(x_i) = -\log_b P(x_i)$

The unit $I(x_i)$ is the binary bunit or $b=2$

$$I(x_i) = -\frac{\log P(x_i)}{\log 2}$$

Q) A source produces one of 4 possible symbols during each interval having probabilities. $P(x_1)=\frac{1}{2}$, $P(x_2)=\frac{1}{4}$, $P(x_3)=\frac{1}{8}=P(x_4)$ obtain the information content of each symbol?

→ $I(x_1) = -\log_2 P(x_1)$

$= -\frac{\log(1/2)}{\log 2} = \underline{1\,bit}$

$I(x_2) = -\log_2 P(x_2)$

$= -\frac{\log(1/4)}{\log 2} = \underline{2\,bits}$

$I(x_3) = -\log_2 P(x_3)$

$= -\frac{\log(1/8)}{\log 2} = \underline{3\,bits} = I(x_4)$

This revels the fact that if probability of occurance is less this information carried is more & vice versa.

Q) In a binary PCM if zero occur with probability 1/4 & '1' occur probability 3/4 then calculate the amount of info carried by each beneath?

→ $I(x_0) = -\log_2 P(x_0)$

$= -\frac{\log (1/4)}{\log 2} = \underline{2\,bits}$

$I(x_1) = -\log_2 P(x_1)$

$I(x_1) = -\frac{\log(3/4)}{\log 2}$

$= \underline{0.415\,bit}$

Entropy (Average information):

In practical communication system we usually transmit long sequences of symbol from a information source thus we are more intrested in Avg info. that a source produces than the info. content of a singal symbol. ~~In order to get the info. content of~~ symbol. The avg. value of $I(x_i)$ over the alphabet of source x of m different symbols is given by

$$H(x) = E[I(x_i)] = \sum_{i=1}^{m} P(x_i) I(x_i)$$

↓

Entropy

$$\boxed{H(x) = -\sum_{i=1}^{m} P(x_i) \log_2 P(x_i) \quad bits/symbol}$$

The quantity H(x) is called the entropy of source x it is the measure of the avg info. content per source symbol.

Information Rate:

The rate at which source 'X' emits symbols is 'r' (sym/sec)

⇒ the information Rate (R) of the source (X) is given by

$\boxed{R = r\,H(X)\ b/s}$ r → rate at which symbol are generated
H(X) → Avg info. / Entropy (b/symb)

Q) Discrete memoryless Source (DMS) X has 4 symbols $x_1, x_2 \ldots x_4$ with probabilities 0.4, 0.3, 0.2, 0.1 resp. calculate H(x). Find the amount of information contained in the messages x_1, x_2, x_1, x_3 & $x_4 x_3 x_3 x_2$ compare with the H(x) obtained in part (i)

→ $H(x) = -\sum_{i=1}^{m} P(x_i) \log(P(x_i))$

$= -\left(0.4 \cdot \frac{\log 0.4}{\log 2} + 0.3 \frac{\log 0.3}{\log 2} + 0.2 \frac{\log 0.2}{\log 2} + 0.1 \frac{\log 0.1}{\log 2}\right)$

$= -(-0.528 - 0.521 - 0.464 - 0.332)$

$= \underline{1.845 \text{ bits/symbol}}$

$I(x_1 x_2 x_1 x_3) = -\log_2 P(x_1 x_2 x_1 x_3)$

$= -\log_2 P(0.0048) = -\frac{\log(0.0048)}{\log 2}$

$= \underline{7.702 \text{ bits}}$

$I(x_4 x_3 x_3 x_2) = -\log_2 P(0.0012) = \underline{9.702 \text{ bits}}$

Q) Discrete source emits symbols once in every millisec, with probabilities 1/2, 1/4, 1/8, 1/16, 1/16 resp. Determine the source entropy and information rate?

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

$= -\left(\frac{1}{2} \frac{\log(1/2)}{\log 2} + \frac{1}{4} \frac{\log(1/4)}{\log 2} + \frac{1}{8} \frac{\log(1/8)}{\log 2} + \frac{1}{8} \frac{\log(1/16)}{\log 2}\right)$

$= -\left(-0.5 + (-0.5) - \frac{3}{8} - 0.5\right)$

$= \underline{1.875 \text{ bits/symbol}}$

$P(X) = P(x_1 x_2 x_3 x_4 x_5)$

$= 8 (8.10^{-3} \times 10^{-3})$

$I(X) = -\log P$

$R = r\, H(x) = 1.875$ kb/s

$R = \frac{1}{10^{-3}} H(x)$

Q) The probability of five possible experiments are given as $P(x_1) = 0.5$, $P(x_2) = 1/4$, $P(x_3) = 1/8$, $P(x_4) = P(x_5) = 1/16$. Determine the entropy and Infor rate if there are 16 outcomes /sec.

→ $H(x) = 1.875$

$R = r\, H(x)$

$= 16 (1.875) = 30$ b/sec

Q) An analog signal band limited to 10K Hz is quantized in 8 levels of a PCM system with probabilities of 1/4, 1/5, 1/5, 1/10, 1/10, 1/20, 1/20, 1/20, respectively, find $H(X)$ & R?

→ By Nyquist criteria

$f_s \geq 2 f_m$

$f_s \geq 20$ KHz

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

$$= -\left(\frac{1}{4} \frac{\log(1/4)}{\log 2} + \frac{2}{5} \frac{\log(1/5)}{\log 2} + \frac{2}{10} \frac{\log(1/10)}{\log 2} + \frac{3}{20} \frac{\log(1/20)}{\log 2} \right.$$

$$= -\left(-0.5 + \frac{2}{5}.928 - 0.66 - 0.648 \right)$$

$H(x) = 2.73$ bits / symbol

$r = 20K$

$R = 20K (2.73)$

$= 54.7$ K bit/sec

Q) A high-resolution black & white tv picture consists of about 2×10^6 picture elements & 16 different brightness level. Pictures are repeated at the rate 32/sec all picture element are assumed to be independent and all levels have equal likelywood of occurrence? calculate avg rate of info. conveyed by this tv picture source?

→ $m = 16$ $\quad p(x_i) = 1/16$

$H(x) = -\sum_{i=1}^{16} \left(\frac{16}{16} \frac{\log(1/16)}{\log 2} \right)$

$H(x) =$ 4 bits /symbols

$R = r\,H(x)$

$= 2\times10^6 \times 32 \times 16$

$= 2$

Q Given a telegraph source having 2 symbols (dot & –) the dot duration is 0.2 sec the – duration is 3 times dot. the $p(\cdot) = 2p(-)$ & the time b/w symbols is 0.2 sec calculate the information rate of telegraph source?

→ $p(\cdot) = 2p(-)$

$p(\cdot) + p(-) = 3p(-) = 1$

from above two equation $p(-) = \frac{1}{3}$ $\quad p(\cdot) = \frac{2}{3}$

$H(x) = -\sum_{i=1}^{2} p(x_i) \log_2 p(x_i)$

$= -\left(\frac{1}{3} \frac{\log(1/3)}{\log 2} + \frac{2}{3} \frac{\log 2/3}{\log 2} \right)$

$= -(-0.52 - 0.389) =$ 0.91 bits /symbol.

$R = r\,H(x)$ $\qquad t_{\cdot} = 0.2 \qquad t_{-} = 0.6s, \; t_{space} = 0.2$

~~T = t. t- ttspace =~~

$T_s = p(\cdot)\, t_{\cdot} + p(-) t_{-} + p(\,)\, t_{space}$

$T_s = 0.533$ s/symbols $\qquad r = \frac{1}{Ts} = 1.875$ symbol/sec

$R = 1.875(0.92) = 1.725$ bits /sec

Discrete memoryless channels (DMC)

Channel representation

$$[x] = \left.\begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix}\right\} x \rightarrow P(y_j/x_i) \rightarrow y \left\{\begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{matrix}\right. [y]$$

A communication channel may be defined as a path or medium through which the symbols flow to the receiver end. DMC is a statistical model with a input X and a o/p Y as shown above. During each unit of the time interval the channel accepts the i/p symbol from [X] and in response it generates an o/p symbol [Y]. This channel is said to be discrete when the alphabets of [X] & [Y] are both finite also it is said to be memoryless & current o/p depends on only current i/p and not on any of the previous i/p. The possible i/p o/p path is indicated along with a conditional probability $P(y_j/x_i)$, where $P(y_j/x_i)$ is the conditional probability of obtaining o/p y_j given that the i/p is x_i. and is called a channel transition probability

Channel matrix:

$$[P(y/x)] = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) & \cdots & P(y_n/x_1) \\ P(y_1/x_2) & P(y_2/x_2) & \cdots & P(y_n/x_2) \\ \vdots & & & \\ P(y_1/x_n) & P(y_2/x_n) & \cdots & P(y_n/x_n) \end{bmatrix}$$

Since each i/p to channel results in some o/p each row of the channel matrix must sum to unity

i.e, $\boxed{\sum_{j=1}^{n} P(y_j/x_i) = 1}$ for all i

Now if the i/p probability $P(x)$ are represented by row matrix then we have $[P(X)] = [P(x_1), P(x_2) \cdots P(x_m)]$ and also the o/p probabilities $[P(Y)] = [P(y_1), P(y_2) \cdots P(y_n)]$

Then $\boxed{[P(Y)] = [P(X)]\,P[Y|X]}$

If $P(X)$ is represented as diagonal matrix, then we have

$$[P(X_d)] = \begin{bmatrix} P(x_1) & 0 & 0 & \cdots & 0 \\ 0 & P(x_2) & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & P(x_m) \end{bmatrix}$$

The elements of

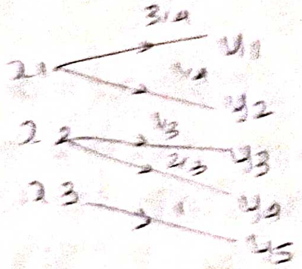
$$\boxed{P[X,Y] = P[X_d]\,P[Y/X]}$$

where i,j elements of matrix $p(x,y)$ has the form $p(x_i, y_j)$. The matrix $p(x,y)$ is known a "joint probability matrix" of i/p x_i and o/p y_j

Types of channel:

Loss less channel:

$$P(Y|X) = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

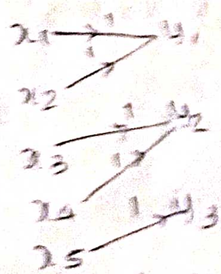


A channel is described by a channel matrix with only one non zero element in each column is called a loss less channel. It can be shown that in loss less channel no source info is lost in transmission.

Deterministic channel:

It is discribed by a channel matrix with only one non zero element in each row is called deterministic channel

$$P(Y|X) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

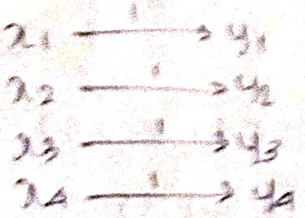


when a given source symbol is sent in deterministic channel it is clear that which o/p symbol will be received

Noise less channel

A channel is called noise less if it is both loss less and deterministic

$$P(Y|X) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



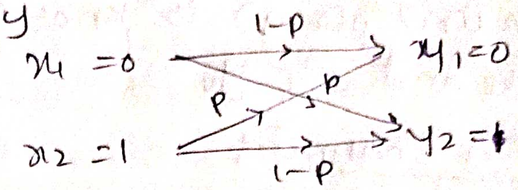
A noiseless channel matrix is shown in figure. the matrix has only one element in each row and each column & this element is unity. Note that the i/p & o/p symbols are same size i.e, $m=n$

Binary symmetric channels:

It is defined by channel diagram

- its channel matrix is given by

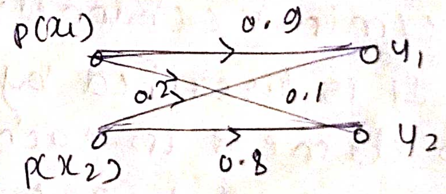
$$[P(y|x)] = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$



It has 2 i/p $x_1 = 0$ and $x_2 = 1$ & 2 o/p $y_1 = 0$, $y_2 = 1$. this channel is symmteric bcz, the probability of receiving a 1 if a zero is sent is same as the probability of reciving a 0 if a one is sent. This common transition probability is denoted by p is shown above.

Q) Given a binary channel shown in fig. find channel matrix of the channel 2) Find probability of y_1 & y_2 when $p(x_1) = p(x_2) = 0.5$, 3) Find the joint probabilities of $P(x_1, y_2)$ & $P(x_2, y_1)$ when $p(x_1) = p(x_2) = 0.5$

→ $P(y|x) = \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \end{bmatrix}$



$$p(y/x) = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$P(y) = P(x)\, P(y/x)$$

$$= [P(x_1) \;\; P(x_2)] \begin{bmatrix} P(y_1/x_1) & P(y_2/x_1) \\ P(y_1/x_2) & P(y_2/x_2) \end{bmatrix}$$

$$= [0.5 \;\; 0.5] \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$P[y] = [0.55 \;\; 0.45]$$

$$\Rightarrow P(y_1) = 0.55 \qquad P(y_2) = 0.45$$

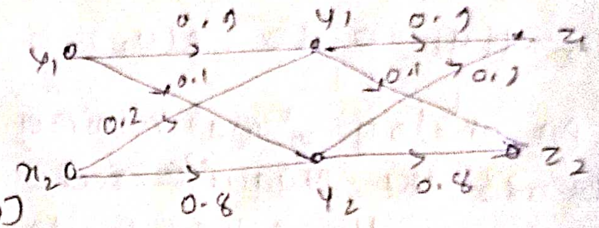
$P(X, Y) = P[X d] P[Y|X]$

$$= \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$= \begin{bmatrix} 0.45 & 0.05 \\ 0.10 & 0.40 \end{bmatrix}$$

$P(x_1, y_2) = 0.05$
$P(x_2, y_1) = 0.10$

Q) Two binary channels are connected in cascade as shown in the fig. find the overall channel matrix of resultant channel. and draw the equivalent channel diagram 2) find $P(z_1)$ & $P(z_2)$ when $P(x_1) = P(x_2) = 0.5$



$P[Y] = [P(x)][P(Y|x)]$

$= [P(y)] [P(Z/y)]$

$P[z] = [P(x)] P[(y|x)] P[(z|y)]$

$$\boxed{P[z/x] = [P(y|x)][P[z|y]]}$$

$$P[Z|x] = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

$$= \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

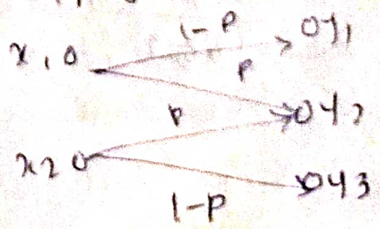
$P[z] = P[x] P[(z|x)]$

$$= [P(x_1) \; P(x_2)] \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

$P(z) = [0.585 \quad 0.415] \Rightarrow P(z_1) = 0.585$
$P(z_2) = 0.415$

Q) A channel has the following channel matrix 1) draw channel diagram 2) The source has equally likely o/p compute the probabilities associated with the channel o/p for p = channel 2

$$\to P(Y|x) = \begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$$



Q) The binary erasure channel has 2 input $x_1 = 0$, $x_2 = 1$ & 3 o/p $y_1 = 0$, $y_2 = e$, $y_3 = 1$, where e indicates an erasure this means that the o/p is in doubt & hence it should be erased 2) $p(y) = [p(x)][p(y/x)]$ substitute values

$$P(Y) = [P(X)][P(Y/X)]$$

$$= [0.5 \quad 0.5] \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

$$= [0.40 \quad 0.20 \quad 0.40]$$

$\rightarrow$ $P(y_1) = 0.40$ $\quad P(y_2) = 0.20$ $\quad P(y_3) = 0.40$

Q) An analog signal having 4 kHz bandwidth is sampled at 1.25 times the Nyquist rate & each sample is quantized into one of equally likely levels assume that successive samples are statistically independent 1) R? 2) can be o/p of the source be transmitted without error over an Additive noise gaussian channel with BW = 10 kHz & SNR = 20 dB. find the SNR required for error free transmission for path 1 4) find the BW required for a AWGN channel for error free transmission of o/p of this source if SNR = 20 dB

$\rightarrow$ $f_m = 4\,kHz$ $\quad f_s = 2f_m$

Nyquist rate $\nearrow$ $\quad r = 2f_m(1.25) = \underline{10\,kHz}$

$r = \underline{10000 \text{ samples/sec}}$

$R = r\,H(x)$

$C = B\log_2\left(1 + \frac{S}{N}\right)$ bits/sec

$C = 10^4 \log_2 [1 + 100]$

$C = 10^4 \frac{\log(101)}{\log 2}$ $\qquad$ $C = \underline{66.6 \times 10^3 \text{ bits/sec}}$

2) The $\frac{S}{N}$ ratio cam be formed by

$$C = 10^4 \log_2 \left(1+\frac{S}{N}\right) \geq \cancel{8} \times 8 \times 10^4$$

$$\log_2 \left[1+\frac{S}{N}\right] > 8$$

$$1+\frac{S}{N} \geq 2^8 = 256 \Rightarrow \frac{S}{N} \geq 255 = \underline{24.1\ dB}$$

$r = r\,H(x)$

$r = 10^4 \times 8 = \underline{80k\ b/s}$

$H(x) = \log_2 (p(x_i)) = \log_2 256 = 8$ bits/samples.

$p(x_i) = \frac{1}{256}$

The required BW can be found by

$$C = B \frac{\log_{10}[1+100]}{\log 2} > 8(10^4) \Rightarrow \underline{\underline{B \geq 1.2\,kHz}}$$

Source coding:

- To conversion of o/p of DMS into a sequence of binary symbols (binary code word) is called source coding.

DMS $x = \{x_1, x_2, x_3\}$ + symbols → Source encoder → binary sequence.

An objective of source coding is to minimize the avg. bit rate required for representation of the source by reducing the information source

Code length and effeciency:

Let x be a DMS with finite entropy $(H(x))$ and an alphabet $x_1, x_2 \ldots x_m$ with corresponding probabilities of occurence $p(x_i)\ \{i = 1, 2 \ldots m\}$, the avg. code word length $\boxed{L = \sum_{i=1}^{m} p(x_i)\, n_i}$ $n_i \rightarrow$ length of code word

The parameter L avg. no. of bits per source symbol used in the source coding process also the code effeciency η is defined as $\boxed{\eta = \frac{L_{min}}{L}}$ $L_{min} \rightarrow$ the minimum possible value of L

When η approaches unity, the code is said to be effecient. The code redundancy, γ is defined as

$$\boxed{\gamma = 1 - \eta}$$

Source coding theorem:

It states that for a DMS (X) with entropy $(H(x))$ the avg. code word length L per symbol bounded by $\boxed{L \geq H(x)}$ and further L can be made as close to $H(x)$ as desired for some suitably choosen code, thus with $L_{min} = H(x)$, the code effeciency can be rewriten as

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

Classification of codes :-

x_i	code 1	code 2	code 3	code 4	code 5	code 6
x_1	00	00	0	0	0	1
x_2	01	01	1	10	01	01
x_3	00	10	00	110	011	001
x_4	10	11	11	111	0111	0001

Fixed length code:

It is one whose code length is fixed. code 1 & code 2 of this table are fixed length with $(L = 2)$.

Variable length codes:

Variable length code is one whose code word length is not fixed. code (3, 4, 5, 6).

Distinct codes:

The code is distinct if each code word is distinguish from other code words. code (2, 3, 4, 5, 6)

Prefix free codes:

A code in which no code word can be formed by adding code wordsymbols to another code word is called a prefix free code. Thus, in a prefix free code, no code word is a prefix of another. code (2, 4, 6)

uniquely decodable codes:

A distinct code uniquely decodable if the original source sequence can be reconstructed perfectly from the endcoded binary sequence. Note that code 3 of the table is not a uniquely decodable code for eg: the binary sequence 1001 may correspond to source sequences x_2, x_3, x_2 and x_2, x_1, x_1, x_2, A sufficient condition to ensure a code is uniquely decodable is that no code word is a prefix of another, thus the prefix free code (2, 4, 6) are uniquely decodable codes.

Instantaneous codes:

An uniquely decodable code is called instantaneous code, the end of any code word is recognizable without examining subsiquent code symbols. so the prefix free codes are sometimes known as instantaneous codes.

optimal codes:

A code is said to be optimal if it is instantaneous and has a minimum avg. L for a given source with a given probability for the source symbols.

Kraft inequality :—

Let X be a DMS with alphabet $x_i = \{i = 1, 2, \dots m\}$ Assume that the length of the assigned binary code word corresponding to $\boxed{x_i \text{ is } n_i}$. A necessary & sufficient condition for exsistnce of a instantaneous binary code is

$$\boxed{K = \sum_{i=1}^{m} 2^{-n_i} \leq 1} \rightarrow \text{kraft inequality}$$

Q) Consider a DMS(X) with symbols $x_i = \{i = 1, 2, 3, 4\}$

i) Show that all the codes except code ~~C~~ satisfy the Kraft inequality

ii) Show that codes (A & D) are uniquely decodable but codes (B ... are not uniquely decodable.

→

$$K = \sum_{i=1}^{m} 2^{-n_i}$$

$$= \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8}$$

$$+ \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{2}$$

$$+ \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = 1$$

x_i	Code A	Code B	Code C	Cod
x_1	00	0	0	0
x_2	01	10	11	100
x_3	10	11	100	110
x_4	11	110	110	111
	2	1223	1233	133
K =	1	1.12	1	0.8

Code A & D are prefix free codes ∴ uniquely decodable.

Entropy Coding :-

The design of a variable length code such that its avg. code word length approaches the entropy of DM is often refered to as entropy coding.

Shanon fano coding:

An efficient code can be obtained by following simple procedure known as shanon fano algorithm

1) List the source symbols in order of decresing probability.

2) partition the set into two set that are as close to equi-probable as possible, and assign 'zero' to upperset and '1' to lower set.

3) Continue this process each time partitioning the sets with as nearly equal probabilities as possible until further partioning is not possible.

Q) A DMS x_i has 6 symbols $x_1, x_2, x_3 \ldots x_6$ with probability $P(x_1) = 0.30$, $P(x_2) = 0.25$, $P(x_3) = 0.20$, $P(x_4) = 0.12$, $P(x_5) = 0.08$, $P(x_6) = 0.05$. Construct a shanon fano code for x, find the efficieny and redundancy of the code?

x_i	$p(x_i)$	step1(.)	step2(.)	step3(.)	step4(.)
x_1	0.30	0	0	0	
x_2	0.25	0	1	1	
x_3	0.20	1	0		
x_4	0.12	1	1	0	
x_5	0.08	1	1	1	0
x_6	0.05	1	1	1	1

	Code	n_i
x_1	00	2
x_2	01	2
x_3	10	2
x_4	110	3
x_5	1110	4
x_6	1111	4

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

$$H(x) = -\sum_{n=1}^{\infty} p(x_i) \log_2 p(x_i)$$

$$H(x) = 0.30 \log_2 0.30 + 0.25 \log_2 0.25 + 0.2 \log_2 0.2 + 0.12 \log_2 0.12 + 0.08 \log_2 0.08 + 0.05 \log_2 0.05$$

$$H(x) = +2.36$$

$$L = \sum_{i=1}^{6} p(x_i) n_i = 0.30(2) + 0.25(2) + 0.20(2) + 0.12(3) + 0.08(4) + 0.05(4)$$

$$= 0.60 + 0.5 + 0.4 + 0.36 + 0.32 + 0.20$$

$$= 1.1 + 0.76 + 0.52$$

$$= 1.86 + 0.52 = 2.38$$

$$\eta = \frac{2.36}{2.38} = 0.991 \times 100\% = 99.16\%$$

Q) A DMS x has 5 symbols $x_1, x_2 \dots x_5$ with $p(x_1) = 0.4$, $p(x_2) = 0.19$, $p(x_3) = 0.16$, $p(x_4) = 0.15$, $p(x_5) = 0.1$, construct a shanon fano code & calculate the effeciency of code

x_i	$p(x_i)$	Step1	Step2	Step3	Code	n_i
x_1	0.4	0	0		00	2
x_2	0.19	0	1		01	2
x_3	0.16	1	0		10	2
x_4	0.15	1	1	0	110	3
x_5	0.1	1	1	1	111	3

$$H(x) = -\sum_{n=1}^{5} p(x_i) \log p(x_i)$$

$$= -(0.4 \log_2 0.4 + 0.19 \log_2 0.19 + 0.16 \log_2 0.16 + 0.15 \log_2 0.15 + 0.1 \log_2 0.1)$$

$$= \underline{2.14}$$

$$L = \sum_{n=1}^{5} p(x_i) n_i$$

$$= 0.4(2) + 0.19(2) + 0.16(2) + 0.15(3) + 0.1(3)$$

$$= 0.8 + 0.38 + 0.32 + 0.45 + 0.3$$

$$= \underline{2.25} \qquad \eta = \frac{2.14}{2.25} = \underline{95.11\%}$$

Huffman Coding:

In general huffman encoding results in an optimum code thus it is the code that has the highest effeciency
1) list the source symbols in order of decresing probability.
2) Combine the probabilities of 2 symbols having the lowest probabilities & reorder the resultant probabilities This step is known as reduction (1). The same procedure is repeated until there are 2 order probabilities remaining.
3) Start encoding with the last reduction which consist of exactly two order probabilities. Assign two to first set & '1' to the lowerset.
4) Now go back & assign 0 & 1 two the second digit for the two probabilities that were combined in previous reduction step.
5) Keep regressing this way until first column is reached

Q)

x_i	$p(x_i)$						Code	n_i
x_1	0.30	0.30	0.30	0.45	0.55	0	00	2
x_2	0.25	0.25	0.25	0.30 0	0.45	1	01	2
x_3	0.20	0.20	0.25 0	0.25 1			11	2
x_4	0.12	0.13 0	0.20 1				101	3
x_5	0.08 0	0.12 1					1000	4
x_6	0.05 1						1001	4

$H(x) = 2.36$, $L = 2.38$

$\eta = 99.16\%$

Q) A DMS X has 5 symbols $x_1 \ldots x_5$, $p(x_1) = 0.4$, $p(x_2) = 0.19$, $(x_3) = 0.16$, $(x_4) = 0.15$, $(x_5) = 0.1$.

→

x_i	$p(x_i)$				Code	n_i
x_1	0.4	0.4	0.4	0.6, 0	1	1
x_2	0.19	0.25	0.35 0	0.4, 1	000	3
x_3	0.16	0.19 0	0.25 1		001	3
x_4	0.15 0	0.16 1			010	3
x_5	0.1 1				011	3

$H(x) = 2.14$

$L = \sum p(x_i) n_i$

$= 1(0.4) + 3(0.19) + 3(0.16) + 3(0.15) + 3(0.1)$

$= 0.4 + 0.57 + 0.48 + 0.45 + 0.3$

$= 2.35$

$\eta = \frac{2.14}{2.2} = 97.3\%$