Information theory and coding

- Information → message
  Information can be analog or discrete

source → source encoder → Noisy chanel modulator amplr → source decoder → destination

- Measure of Information

Information is transmitted from source which is digital or analogs. Let the discrete source descending symbols s_1, s_2 with prob. of occurance p_1, p_2. Out of these symbol if we choose s_k symbol with prob. p_k.

When prob. of occurence is low that is uncertainty is high then it contains information, whereas, if uncertainty is high then there is less amount of information. So the measure of inf. was quantatively defined by shannon

$$I_k = \log_2\left(\frac{1}{p}\right)$$

↓

symbol s_k's probability p_k

I_k is self information of message s_k and b is base of logarithm. Since $0 \leq p_k \leq 0-1$, $I \geq 0$. Information is always non-negetive

The amt of information in a message depends only on its probability of occ. rather than its actual content or its physical interpretation

Consider example of tossing a fair coin. A binary source is producing outcome 0 to tail and 1 for head

$P(1) = 0.5$, $P(0) = 0.5$

$$I(H) = -\log_2(0.5) = \log(2)$$

$$= -\frac{\log(0.5)}{\log(2)}$$

$= 1$ bit

For a discrete memoryless source, each massage is independant of previous massage

- Properties of Information

1) More uncertainty of massage, more the information required

2) If reciever knows massage transmitted, the information 0

3) Total information for n symbols $S_0, S_1, S_2 - S_n$ with prob. P_1, P_2 then total inf. $\rightarrow I_T = \log_b\left(\frac{1}{P_0}\right) + \log_b\left(\frac{1}{P_1}\right) + \log_b\left(\frac{1}{P_2}\right) + \ldots$

4) If there are $M = 2^n$ equily likely messages, then amt of information carried by each message will be N bits

ex1- The binary symbols 0 and 1 are transmitted with probabilities $P(0) = 1/4$, $P(1) = 3/4$ Find the corresponding self information

$I_0 = \log_2(4)$
$= 2$ bits

$I_1 = \frac{\log 4}{\log 2} \rightarrow \log_2\left(\frac{4}{3}\right) = \frac{0.125}{0.301}$

$= 0.415$ bits

Highly probable unit gives less information

- Ex-2- Prove the following statement. If the receiver knows the message transmitted, the amt of information is 0

$P(k) = 1$

- Entropy:-

$I_{Total} = L\left\{ P_0 \times \log_2\left(\frac{1}{P_0}\right) + P_1 \log_2\left(\frac{1}{P_1}\right) + P_2 \log_2\left(\frac{1}{P_2}\right) \right\}$

$\frac{I_{Total}}{L} = \left\{ P_0 \log_2\left(\frac{1}{P_0}\right) + P_1 \log_2\left(\frac{1}{P_1}\right) + P_2 \log_2\left(\frac{1}{P_2}\right) \right\}$

$H(S)$ - Entropy = avg information

Q: A discrete memoryless source produces 4 symbols x_1, x_2, x_3, x_4 with probability $P(x_1) = 0.5$, $P(x_2) = 0.2$, $P(x_3) = 0.2$, $P(x_4) = 0.1$
$H(x)$. Find entropy or average information
Also obtain information contained $X_4 X_3 X_1 X_1$

$\rightarrow$ Entropy = $\sum_{i=1}^{4} P(x_i) \log_2 \left(\frac{1}{Px_i}\right)$

$Px_1 \log\left(\frac{1}{Px_1}\right) + P_2 \log\left(\frac{1}{Px_2}\right) + P_3 \log\left(\frac{1}{Px_3}\right) + P_4 \log\left(\frac{1}{Px_4}\right)$

$0.5 \log\left(\frac{1}{0.5}\right) + 0.2 \log_2\left(\frac{1}{0.2}\right) + 0.2 \log_2\left(\frac{1}{0.2}\right) + 0.1 \log_2\left(\frac{1}{0.1}\right)$

$= 0.5 + 0.46 + 0.464 + 0.332$

$H(x) = 1.76$ bits/symbol

b) $P(x_4 x_3 x_1 x_1) = 0.1 \times 0.2 \times (0.5)^2$

$= 5 \times 10^{-3}$

$I(x_4 x_3 x_1 x_1) = \log_2\left(\frac{1}{P(x_4 x_3 x_1 x_1)}\right) = -\log_2(5 \times 10^{-3})$

$= 7.643$ bits/symbol

Q. Find the relationship between Harley, nats, bits

$\rightarrow$ $I = \log_{10}\left(\frac{1}{P}\right)$ Harlley

$I = \log_e\left(\frac{1}{P}\right)$ nat

$I = \log_2\left(\frac{1}{P}\right)$ bit

$$1 H = \frac{I}{\log_{10}\left(\frac{1}{P}\right)}$$

$$\left[\log_a b = \frac{1}{\log_b a}\right]$$

$$1 H = \frac{\log_e\left(\frac{1}{P}\right)}{\log_{10}\left(\frac{1}{P}\right)} \text{ nats}$$

$$\frac{\log_P 10}{\log_P e} = 2.303 \text{ nats}$$

- $$1 H = \frac{\log_2\left(\frac{1}{P}\right)}{\log_{10}\left(\frac{1}{P}\right)} \text{ bits}$$

$$\frac{\log_P 10}{\log_P 2} = 3.32 \text{ bits}$$

- 1 nat = $\log_2 e$ bits

  1 nat = 1.443 bits

Q An analog

- Information Rate:-

let us suppose that symbols are emitted by source at a fixed time rate "rs" symbol/ sec. the avg source rate "Rs" in bits/ sec are defined as product of (avg. information content per symbol) X (message symbol rate, "rs")

$R_s = H(s) \times r_s$ bits / sec

$R_s = \frac{H}{T}$

where T is time required to send single message

Ex- An event has 4 possible outcomes with prob of occurrence $P_1 = \frac{1}{4}$, $P_2 = \frac{1}{2}$, $P_3 = \frac{1}{8}$, $P_4 = \frac{1}{8}$. Determine the entropy of the system. Also obtain the rate of information if there are 8 outcomes per second

→ $H = \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{2}\log 2 + \frac{1}{8}\log 8$

$\frac{1}{2} + \frac{3}{8} + \frac{1}{2} + \frac{3}{8}$

$= 1 + \frac{3}{4}$

$= \frac{7}{4}$

$= 1.75$ bits / message

$R_s = H(S) \times r_s$

1.75 bits / message × 8 outcome / sec

$= 14$ bits / sec

- An output of information source consist of 150 symbol, 32 of which occur with prob. of $\frac{1}{64}$ and remaining 118 occur with prob of $\frac{1}{236}$. The source emits 2000 symbols per sec. assuming that symbols are chosen independently. Find avg information rate for the source

→ $L = 150$

$$H(x) = \sum_{i=1}^{32} P_{xi} \log\left(\frac{1}{P_{xi}}\right) + \sum_{33}^{150} P_{xi} \log_2\left(\frac{1}{P_{xi}}\right)$$

$$= 3 + 3.941$$

$$= 6.94 \text{ bits/message}$$

↓
symbol

$$R = H \times r$$

$$= 6.94 \times 2000$$

$$= 13880 \text{ bits/sec}$$

↓
symbol

$$= 13880 \text{ bits/symbol}$$

- Let us take first letter A to H, D find code length

letters	code used	letters	code used
A	000	E	100
B	001	F	101
C	010	G	110
D	011	H	111

- Variable code length (VLC) 1

letters	code used
A	00
B	010
C	011
D	100
E	101
F	110
G	1110
H	1111

- Variable code length (VLC) 2

letter	code used
A	0
B	1
C	00
D	01
E	10
F	11
G	000
H	111

Encode series of letters "ABADCAB"

Message - A B A D C A B
fixed - 000 001 000 011 010 000 001
VLC1 - 0001000100011000010
VLC2 - 010010001

Total bits - 21 + 18 + 9
= 48

Regroup the bits of VLC2 in different manner

VLC2 - [0] [10] [0] [1] [0] [0] [01] ... (Grouping)

AEABAAD - decoded message 1

Regroup single single bits

VL2 - [0] [1] [0] [0] [1] [0] [0] [0] [1]
ABAABAAAB - decoded message 2

- Prefix Code:-

Prefix codes are the codes in which no code word forms the prefix of any other code

- Instantaneous code:-

As soon as sequence of bits corresponding to any one of the corresponding code word is detected, we can declare that symbol is decoded. This code is called instantaneous code

There is no decoding delay in instantaneous codes

Q.1 Consider the following where source of size 4 has been encoded in binary codes with 0 and 1. Identify different codes

x_i	code 1	code 2	code 3	code 4	code 5	co
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	11	11	11	111	0111	0001

- Types of code

1) Fixed code / length - code 1, code 2
2) Variable length - code 3, code 4, code 5, code 6
3) Instantaneous code - } 2, 4, 6
4) Prefix code - }
5) Uniquely decodable -
6) Distinct Codes -

Q.2 Following example has 2 binary code for 4 symbol each having probability 0.5, 0.25, 0.125, 0.125 resp Compare their efficiency

- codelength - No of bits for encoding. Avg code length is - . Average no of bits per source letter or symbol $= \bar{L} = \sum_{i=1}^{N} n_i P(x_i)$

- code efficiency

$$\eta = \frac{H(s)}{L} \text{ - entropy}$$

- γ = code redundancy

$$\gamma = 1 - \eta$$

code 1	n_i	code 2	n_i
00	2	0	1
01	2	10	2
10	2	110	3
11	2	111	3

x_i	$P(x_i)$
x_1	0.5
x_2	0.25
x_3	0.125
x_4	0.125

- Code 1

Avg length

$$L_{code 1} = n_1P_1 + n_2P_2 + n_3P_3 + n_4P_4$$

$$H(x)_{code 1} = \sum_{i=1}^{4} P_{xi} \log_2\left(\frac{1}{P_{xi}}\right) = 1.75 \text{ bits/symbol}$$

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

$L = 2$ bits/symbol

$$\eta = \frac{1.75}{2} = 0.875$$

$$\eta\% = 87.5\%$$

- Optimum Code - If it is instantaneous and has minimum average length $\bar{L}$ for given source for with the particular prob assignment for source symbol

- Entropy code :- When a variable length code is designed such that its average code per length $\bar{L}$ approaches the entropy of DMS (discrete memory less source)

- H(x) for VLC2 (question continued)

$$L = n_1P_1 + n_2P_2 + n_3P_3 + n_4P_4$$

$$0.5 \times 1 + 0.25 \times 2 + 0.125 \times 3 \times 2$$

$$L = 1.75$$

$$H = 1.75$$

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

$$\eta = \frac{1.75}{1.75}$$

$$\eta = 1 \; , \; \eta\% = 100\%$$

Q. A binary source is emitting an independent source of 0 and 1 with prob $P(0) = P$, $P(1) = 1-P$. Plot the entropy of source VS P

→ $H(x) = \sum_{i=1}^{2} P_{xi} \log_2 \left(\frac{1}{P_{xi}}\right)$

$$H(u) = P \log_2\left(\frac{1}{P}\right) + (1-P)\log_2\left(\frac{1}{1-P}\right)$$

H.W Make graph

- Shannon Fano coding

1) It is a source coding technique for constructing prefix code based on set of symbols and based on probabilities. This is sub-optimal code as it does not achieve lowest possible codeword length.

2) It produces fairly efficient variable length coding

3) It is used in IMPLODE compression method which is part of zip file format

- Steps of algorithm

1) Arrange the source ~~so~~ symbol in order of decreasing order. The symbols which equal probabilities can be listed in any arbitary order

2) Divide the set into 2, such that the sum of the probabilities in each set is same or nearly same.

3) Assign 0 to the upper set and 1 to the lower set

4) Repeat steps 2 and 3 until each subset contains a single symbol

Qn1- A DMS has five symbols x_i, with their probabilities as follows. Construct the Shannon-Fano code for X and calculate efficiency of code

x_i	P_{x_i}
x_1	0.4
x_2	0.17
x_2	0.18
x_3	
x_4	0.1
x_5	0.15

→ Arranging

Take the group of 2 such that sum were no

x_i	P_{x_i}	column 1	column 2
x_1	0.4	0	0
x_3	0.18	0	1
x_2	0.17	1	0
x_5	0.15	1	1
x_4	0.1	1	1

column 3	code	Codeword
0	00	2
1	01	2
0	10	2
0	110	3
1	111	3

$L = P \times n$

$= 0.4 \times 2 + 0.18 \times 2 + 0.17 \times 2 + 0.15 \times 3 + 0.1 \times 3$

$0.8 + 0.36 + 0.34 + 0.45 + 0.3 = 2.25$ bits/symbol

$H(x) = 0.4 \log\left(\frac{1}{0.4}\right) + 0.18 \log\left(\frac{1}{0.18}\right) + 0.17 \log\left(\frac{1}{0.17}\right) + 0.15 \log \frac{1}{0.15}$

$+ 0.3 \log\left(\frac{1}{0.3}\right) = 2.149$ bits/symbol

$\eta = \frac{2.14}{2.25} = 0.955$

$\therefore \eta\% = 95.5\%$

another way

x_i	P_{x_i}	column1	column2	column3	code	Length
x_1	0.4	0	0	0	00	2
x_2	0.18	1	0	0	100	3
x_3	0.17	1	0	1	101	3
x_4	0.15	1	1	0	110	3
x_5	0.1	1	1	1	111	3

$L = 2.2$ bits/symbol

$H(x) = 2.149$ bits/symbol

$\eta = \frac{2.14}{2.2} = 0.9727$

$\therefore \eta\% = 0.97 \times 100$

$= 97.27\%$ - code efficiency

- Haffmaan coding

Haffmaan coding produces prefix code that always achieve the lowest possible avg code word length. It is an optimal code which has highest efficiency or lowest redundancy. Also known as optimal code and redundancy code

- Procedure for Haffmaan coding :-

1) List the source code in decreasing prob. The symbols with equal prob can be arranged in any arbitary order

2) Combine prob. of symbol having the smallest prob. Now reorder the resultant prob. This process is called reduction 1. The same process is repeated until there are 2 ordered prob. remaining. The final step is called as last reduction

3) Start encoding with the last reduction. Assign 0 as the first digit in the code words for all the source symbols, associated with the first prob of the last reduction. Then assign 1 to the second probability

4) Now go back to the previous reduction step. Assign 0 and 1 to the 2nd digit for the 2 prob that was combined in this reduction step. retaining all assignments made in step 3

5) Repeat step 4 until first column is reached

Ex: Construct Huffman code for the following

x_i	P_{x_i}
x_1	0.4
x_2	0.18
x_3	0.17
x_4	0.15
x_5	0.10

0.15 and 0.10: combine these 2, 0.25, Now arrange in decreasing order again

0.4
0.25
0.18
0.17

Now take 0.18 and 0.17.

combined = 0.35

∴ Now decreasing order - 0.4
0.35
0.25

again do the same process

new combined = 0.6 - reduction 2

∴ Now decreasing order = 0.6
0.4

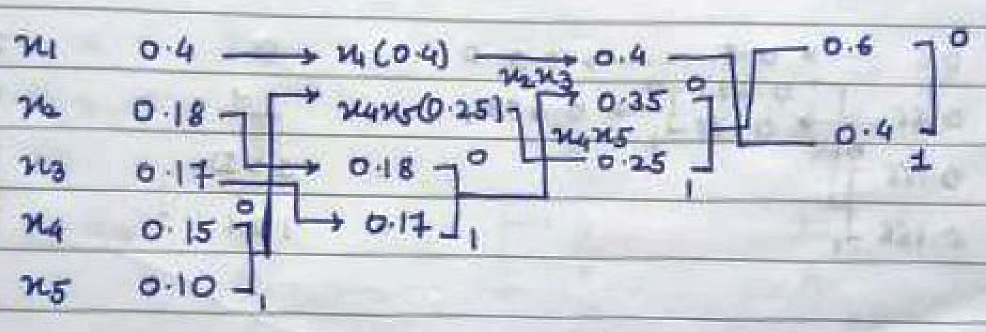
At last 2 probabilities will come - one highest and one lowest

- Kraft's inequality condition

Kraft's inequality is a necessary and sufficient condition to prove existence of prefix code (uniquely decodable code) given n symbol and l_i are no of bits used to represent a symbol for all $i = 1$ to n

$$L = \sum_{i=1}^{n} 2^{-l_i} \le 1$$

- Corrected Method (Huffman code)



x_1 - 1

x_2 - 000

x_3 - 001

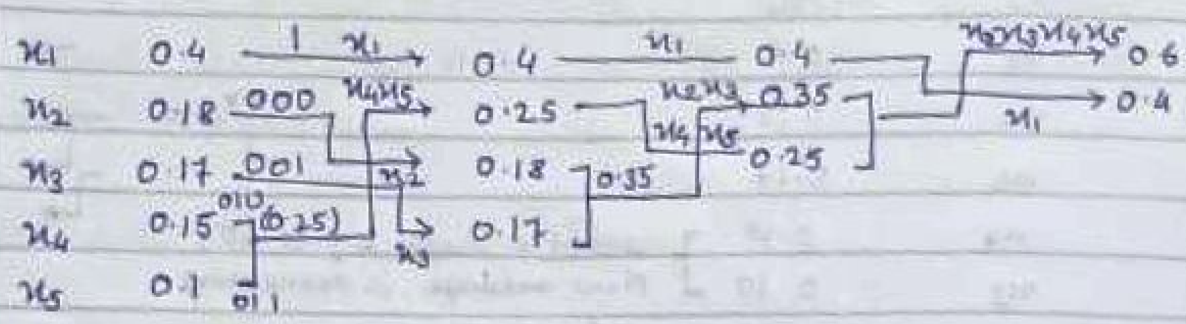
x_4 - 010

x_5 - 011

- Construct Huffman code for the following symbols given

Huffman coding

x_i P_{x_i}



x_i	P_{x_i}	code	code word length
x_1	0.4	1	1
x_2	0.18	000	3
x_3	0.17	001	3
x_4	0.15	010	3
x_5	0.1	011	3

$$H(x) = \sum_{i=1}^{5} P_{x_i} \log_2\left(\frac{1}{P_{x_i}}\right)$$

$$0.4\log_2\left(\frac{1}{0.4}\right) + 0.18\log_2\left(\frac{1}{0.18}\right) + 0.17\log_2\left(\frac{1}{0.17}\right)$$

$$+0.15\log_2\left(\frac{1}{0.15}\right) + 0.11\log_2\left(\frac{1}{0.1}\right)$$

$$H(x) = 2.151 \text{ bits/symbol}$$

$$\bar{L}_{avg} = \sum_{i=1}^{5} n_i P_{x_i} = 2.2$$

$$\eta = \frac{2.151}{2.2} = 0.977 = 97.7\%$$

$\therefore$ The efficiency is 97.7%.

• "SAHARA" - Encode with haffmaan code

$P(S) = \frac{1}{6}$

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

$P(H) = \frac{1}{6}$

$P(R) = \frac{1}{6}$

x_i	P_{x_i}			code	codeword length
x_1 (A)	0.5	→ 0.5	→ 0.5 0	0	1
H	0.166	0.333 0	→ 0.499 1	11	2
R	0.166 0	0.166 1		100	3
S	0.166 1			101	3

$$\bar{L} = 0.5 \times 1 + 0.166 \times 2 + 0.166 \times 6$$

$$0.996 + 0.5 + 0.332$$

$$= 1.828 \text{ bits/symbol}$$

$$H = 1.055 \text{ bits/symbol}$$

$$\eta = \frac{1.055}{1.828}$$

$$= 0.5771$$

$$\eta\% = 57.71\%$$

$\therefore$ The efficiency is 57.71%.

Corrected solution

$H = 1.79$ bits/symbol
$L = 1.83$ bits/symbol

$$\eta = \frac{H}{L} = \frac{1.79}{1.83}$$

$$\eta\% = 97.8\%$$

Ex- A discrete memory less source has 7 signals

x_i	P_{x_i}					
x_1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	......	$\frac{1}{2}$	0
x_2	$\frac{1}{2^2}$	$\frac{1}{4}$	$\frac{1}{4}$	......	$\frac{1}{2}$	1
x_3	$\frac{1}{2^3}$	$\frac{1}{8}$	$\frac{1}{8}$	......		
x_4	$\frac{1}{2^4}$	$\frac{1}{16}$	$\frac{1}{16}$	...		
x_5	$\frac{1}{2^5}$	$\frac{1}{32}$	$\frac{1}{16}$	.....		
x_6	$\frac{1}{2^6}$	$\frac{1}{32}$				
x_7	$\frac{1}{2^{7-1}} = \frac{1}{2^6}$					

Construct huffman code and calculate efficiency of code and comment about kraft's inequality condition

* Here self information of x_1 is 1, x_2 is 2 ....
x_6 is 6
∴ The information (self) is 1 so bits transfer will be 1
∴ l_i (codeword length) = information transfer

KIT tells us about the following

x_i	Px_i	code	codeword length
x_1	$1/2$	0	1
x_2	$1/2^2$	10	2
x_3	$1/2^3$	110	3
x_4	$1/2^4$	1110	4
x_5	$1/2^5$	11110	5
x_6	$1/2^6$	111110	6
x_7	$1/2^6$	1111110	$7-1=6$

$\bar{L} = 1.968$

$\uparrow \sum_{i=1}^{7} n_i P_i$

$\bar{H} = \sum_{i=1}^{7} P_i \log\left(\frac{1}{P_i}\right)$

$= 1.968$

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

$\eta = \frac{1.968}{1.968}$

$\eta = 1$

$\eta\% = 100\%$

$\therefore$ The efficiency is 100%.

- Key point

- Huffman codes are only optimal if the probabilities of symbol are -ve power of 2. This is because all the prefix codes were at bit level

- Arithmatic coding doesn't have this restriction. It work by representing file to be encoded by an interval of real no's between 0 and 1, successive symbol in the ~~wi~~ message reduce this interval a advanced width: the probability of that symbol. The more likely symbols reduce range by less and thus add fewer no of bits to the message.

- Encoding steps:

1) To code symbol 's' when symbols are numbered from 1 to n and symbol i has probability $Pr(i)$.

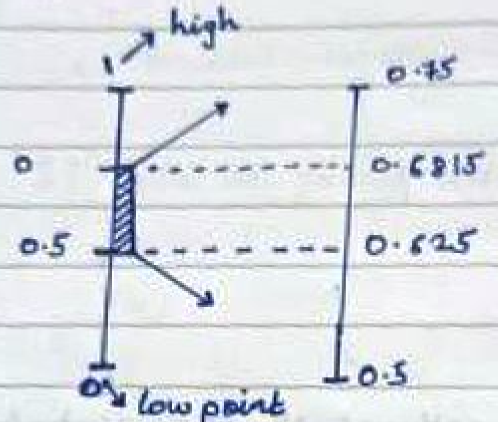
$$\text{low-bound} = \sum_{i=1}^{s-1} Pr[i] \quad \text{range} = \text{high} - \text{low}$$

$$(\text{low}) = \text{low}\ (\text{range} + \text{low bound})$$

$$\text{high-bound} = \sum_{i=1}^{s} Pr[i] \quad , \text{high} = \text{low}\ (\text{range} + \text{high bound})$$

Eg:- Consider symbols 'A', 'B', 'C' with probabilities 0.5, 0.25 and 0.25. Encode given input symbol stream BAC using Arithmetic code

→



For B

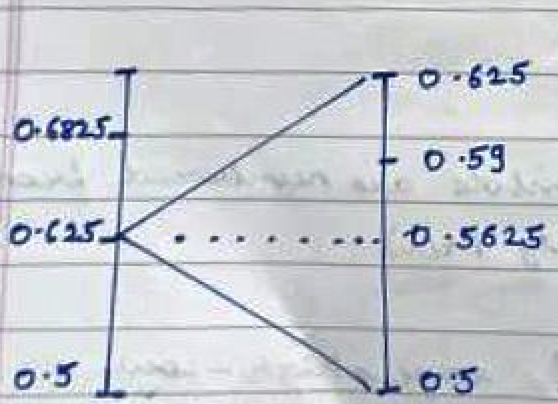
$$0.5 + (0.75 - 0.5) \times 0.5 = 0.625 \quad \dots \text{1st step}$$

(0.5 → low, 0.75 → high, 0.5 → low, 0.5 → Prob.)

$$= 0.625 + (0.75 - 0.5) \times 0.25 = 0.6875 \quad \dots \text{2nd step}$$

$$0.6875 + (0.75 - 0.5) \times 0.25 = 0.75 \quad \dots \text{3rd step}$$

Here we take with each alphabet's probability one by one and the output / answer of first step is the first value of the second step



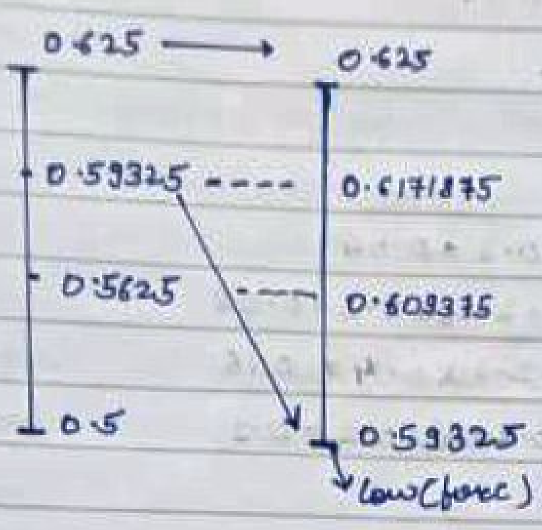
For A

$0.5 + (0.625 - 0.5) \times 0.5 = 0.5625$

$0.5625 + (0.625 - 0.5) \times 0.25 = 0.59375$

$0.59375 + (0.625 - 0.5) \times 0.25 = 0.625$

Now after doing for first letter observe the answers/output reading and according to make diagram and decide high and low values for next letter



For C

$0.59375 + (0.625 - 0.59375) \times 0.5$
$= 0.609375$

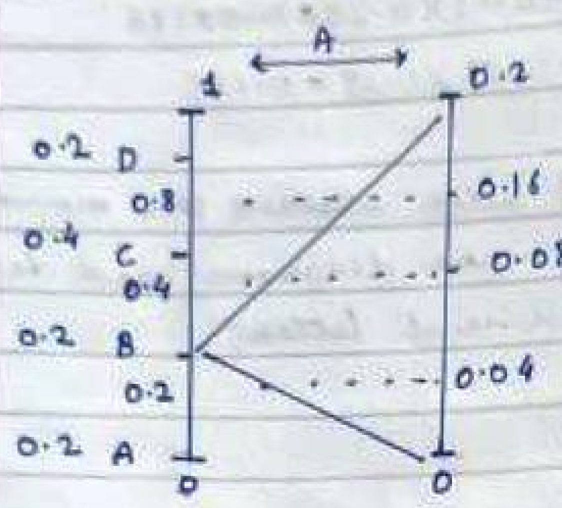
$0.609375 + (0.625 - 0.59375) \times 0.25$
$= 0.6171875$

$0.6171875 + (0.625 - 0.59575) \times 0.25$
$= 0.625$

The final answer lower value of last letter

$\therefore$ Final Answer = "0.59325"

Q Four symbols A, B, C, D with probabilities 0.2, 0.2, 0.4 and 0.2. Transmitted message is ABC



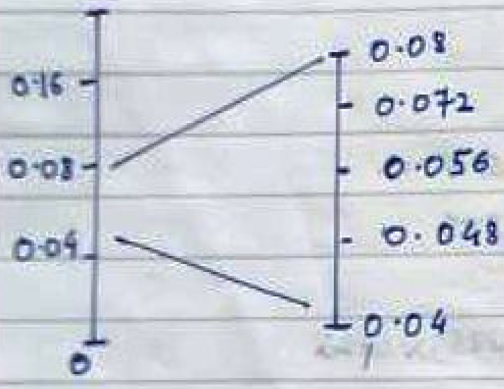
For A

$0 + (0.2-0) \times 0.2 = 0.04$

$0.04 + (0.2-0) \times 0.2 = 0.08$

$0.08 + (0.2-0) \times 0.4 = 0.16$

$0.16 + (0.2-0) \times 0.2 = 0.2$

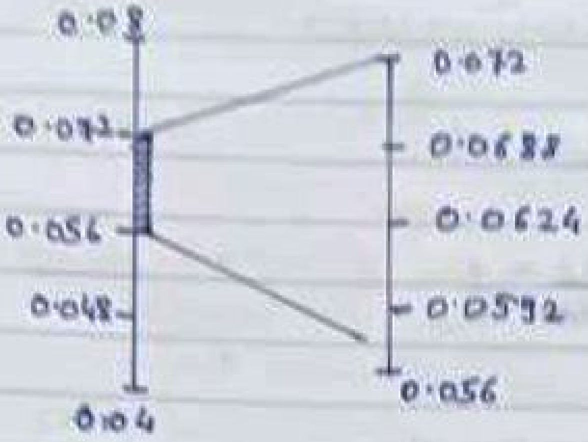


For B

$0.04 + (0.08-0.04) \times 0.2 = 0.048$

$0.048 + (0.08-0.04) \times 0.2 = 0.056$

$0.056 + (0.08-0.04) \times 0.4 = 0.072$



For C

$0.056 + (0.072 - 0.056) \times 0.2 = 0.0592$

$0.0592 + (0.072 - 0.056) \times 0.2 = 0.0624$

$0.0624 + (0.072 - 0.056) \times 0.4 = 0.0688$

The final Answer is low value of last letter

Final Answer = 0.056

Q COFFEE

C = 1/6, O = 1/6, F = 2/6, , E = 2/6

- Solve this by the above method. We can take letters in any order

- Runlength coding

Random bits - 111111111111 0000000111

make groups

12 (1's), 7 (0's), 3 (1's)

represent no in bits

(1100, 1), (111, 0), (11, 1) → 12 bits

original bits = 22 bits

$$\text{Compression} = \frac{12}{22} = \frac{6}{11}$$

- Lempel - Ziv - Walsh (LZW coding)

- Dictionary based coding algorithm

Ex Encode using LZW - "DRESSES"

original dictionary

symbol	code
D	0
R	1
E	2
S	3

Input	current string	seen before	Encoded o/p	New dict entry/index
D	D	Yes	—	—
D|R	DR	No	0	DR|4
DR|E	RE	No	0 1	RE|5
DRE|S	ES	No	0,1, 2	ES|6
DRES|S	SS	No	0, 1, 2, 3	SS|7
DRESS|E	SE	No	0,1,2,3,3	SE|8
DRESSE|S	ES	Yes	—	—
ES	ES	Yes	—	—
	ES	Yes	0,1,2,3,3,6	—

Final code - 0,1, 2, 3, 3, 6

Corrected answer (DADA-DAADAA)

original dictionary

Symbol	code
D	0
A	1
-	2

Input	current string	seen before	encoded O/P	New dictionary entry/index
D	D	Yes	—	—
D|A	DA	No	0	DA/3
DA/D	AD	No	0,1	AD/4
DAD|A	DA	Yes	—	—
DADA|-	DA-	No	0,1,3	DA-/5
DADA-/D	-D	No	0,1,3,2	-D/6
DADA-D/A	DA	Yes	—	—
DADA-DA/A	DAA	No	0,1,3,2,3	DAA/7
DADA-DAA|D	AD	Yes	—	—
DADA-DAAD/D	ADD	No	0,1,3,2,3,4	ADD/8
DADA-DAADD/A	DA	Yes	—	—
—	DA	—	0,1,3,2,3,4,3	—

$\boxed{\text{Final Answer - 0,1,3,2,3,4,3}}$

• Flow chart


Start
Input data
Trace each element and set initial library set last data as empty
Get i/p data
set current input + data (last) as string
string in library
No
Save current string in library
Yes
No encoded output no dictionary entry
Encoded o/p to last data location
set current i/p as last data
Any elements left
No
Yes
end
Set last string as last data

Ex1- Throw a die and obs no of dots appearing on the front face

i) construct sample space

→ $S = \{1, 2, 3, 4, 5, 6\}$

ii) let A be an event that odd no. occurs

→ $\{1, 3, 5\} = S\ (A)$

iii) Let B be an event that even no. occurs

→ $S = \{2, 4, 6\}\ (B)$

iv) let C be an event that prime no occurs

→ $S = \{2, 3, 5\}$

v) find the event that an even or prime no occurs

→ $S = \{2, 3, 4, 5, 6\}$

vi) list the outcomes for an event that even prime no occurs

→ $B \cap C = \{2\}$

vii) find the event that prime does not occur

→ $C = \{1, 4, 6\}$

viii) find the event that seven dots appear on top face

→ $S = \{\phi\}$

ix) find the event that even and odd no occurs simultaneously

→ $A \cap B = \phi$

Ex-2-

	employed	Unemployed	Total
Male	250	50	300
Female	150	100	250
Total	400	150	550

a) If a person is male. What is the prob that he is unemployed

- $0 \leq P(A) \leq 1$
- $P(\bar{A}) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B)$
- $P(\phi) = 0$
- $P(A \cup B) = P(A) + P(B) + P(A \cap B)$
- $P(B/A) = P(A \cap B) / P(A)$

$\because \left[P(U/M) = \frac{P(U \cap M)}{P(M)} = \frac{50}{300} = \frac{1}{6} \right.$

→ a) 50/300 = 1/6 → probability

- Conditional Probability - cond^n prob of an event A given that B has happened, it is defined as

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad ; \; P(B) \geq 0$$

$\because P(A \cap B) \rightarrow$ Joint probability of A and B

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \; ; \quad P(A) > 0$$

$$P(A \cap B) = P(A/B) \cdot P(B) = P(B/A) \cdot P(A)$$

b) If a person is female, what is prob that she is employed

$$P(E/F) = \frac{P(E \cap F)}{P(F)}$$

$$= \frac{150}{250}$$

c) If a person is employed, what is probability he is male

$$P(M|E) = \frac{P(M \cap E)}{P(E)}$$

- Independent event :-

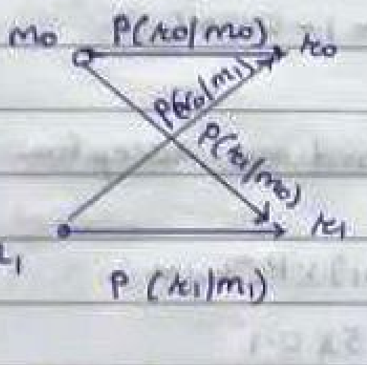
Two events A and B are statistically independent if $P(A|B) = P(A)$ and $P(B|A) = P(B)$

i.e. occurrence or non-occurrence B or A are said to be independent

For independent event - $P(A \cap B) = P(A) \cdot P(B)$

[as $P(A|B) = P(A)$ for independent event]

- In a binary communication system either 0 or 1 is transmitted. Due to channel noise, a one can be received as zero and vice-versa. Let m_0 and m_1 denote event of transmitting 0 and 1 respectively. Let r_0 and r_1 denote event of receiving 0 and 1 respectively



Given $P(m_0) = 0.5$

$P(r_1/m_0) = 0.2$

$P(r_0/m_1) = 0.1$

Find

a) $P(r_0)$, $P(r_1)$

b) If a 1 was recieved, what is probability that a 1 was sent?

c) If a 0 was recieved, what is probability that a 0 was sent

$P(m_0/r_0) = ?$

d) Calculate probability of error

e) Calculate probability of error of Pc that transmitted signal correctly reaches to reciever

→ Sol^n:- $P(m_0) = 0.5$

$P(r_1|m_0) = 0.2$

$P(r_0/m_1) = 0.1$

$P(m_1) = 1 - P(m_0) = 0.5$

$P(r_0|m_0) + P(r_1|m_0) = 1$

$\therefore P(r_0|m_0) = 1 - P(r_1|m_0)$

$1 - 0.2 = 0.8$

$P(r_0|m_1) + P(r_1|m_1) = 1$

$P(r_1|m_1) = 1 - 0.1 = 0.9$

Multiple transmission and one reception

a) $P(r_0) = P(m_0) \times P(r_0|m_0) + P(m_1) \times P(r_0|m_1)$

$0.5 \times 0.8 + 0.5 \times 0.1$

$= 0.45$

$P(x_1) = P(x_1|m_0) \times P(m_0) + P(x_1|m_1) \times P(m_1)$

$0.2 \times 0.5 + 0.9 \times 0.5$

$= 0.55$

b) $P(m_1|x_1) = \frac{P(x_1|m_1) \cdot P(m_1)}{P(x_1)}$

$= \frac{0.9 \times 0.5}{0.55}$

$= 0.818$

c) $P(m_0|x_0) = \frac{P(x_0|m_0) \times P(m_0)}{P(x_0)}$

$= \frac{0.8 \times 0.5}{0.45}$

$= 0.889$

- Total probability

Let A_1, A_2 be usually enclusive $[A_i \cap A_j = \phi]$. Now B is any event in S

Then $P(B) = \sum_{i=1}^{n} P(B \cap A)$

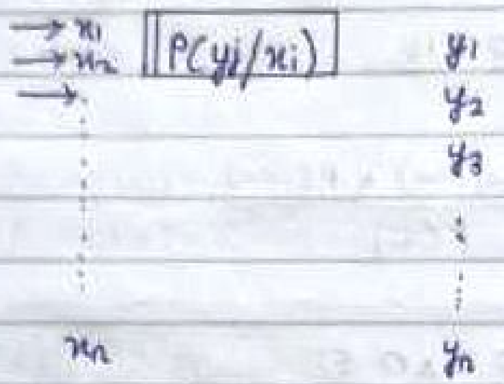
$= \sum_{i=1}^{n} P(B|A_i) P(A_i)$

$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{\sum_{i=1}^{n} P(B|A_i) P(A_i)}$

Channel Model

A communication channel is a transmission path through which symbols flow through the reciever

The - - channel is discrete because alphabets X and Y both are finite. Hence channel is discrete. since present o/p of channel solely depends on the current i/p and not on input of previous stages



X = input
Y = output

$P(y_j/x_i) \longrightarrow$ channel transaction probability

Case 1

$$[P(Y|X)] = \begin{bmatrix} P(y_1|x_1) & P(y_2|x_1) & P(y_3|x_1) & \cdots & P(y_n|x_1) \\ P(y_1|x_2) & P(y_2|x_2) & P(y_3|x_2) & \cdots & P(y_n|x_2) \\ \vdots & \vdots & \vdots & & \vdots \\ P(y_1|x_m) & P(y_2|x_m) & P(y_3|x_m) & \cdots & P(y_n|x_m) \end{bmatrix}$$

1) Each row contains conditional prob of all O/P symbols

$$\sum_{j=1}^{n} P(y_j|x_i) = 1 \quad \text{for all } i$$

2) The input probabilities $P(x)$ are given by a row matrix

3) $\boxed{[P(Y)] = P(Y_1) + P(Y_2) \ldots \; P(Y_n)]}$

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

If $[P(X)]$ is represented as diagonal matrix

$$[P(X)]_d = \begin{bmatrix} P(x_1) & 0 & 0 & \ldots & 0 \\ 0 & P(x_2) & 0 & \ldots & \\ 0 & \vdots & P(x_3) & \ldots & \\ \vdots & & \vdots & & \\ 0 & 0 & 0 & \ldots & P(x_m) \end{bmatrix}$$

$$\boxed{\text{Then} \quad [P(X,Y)] = [P(x)]_d \cdot [P(y|x)]}$$

$[P(X,Y)]$ matrix is of form $P(x_i, y_j)$

↓

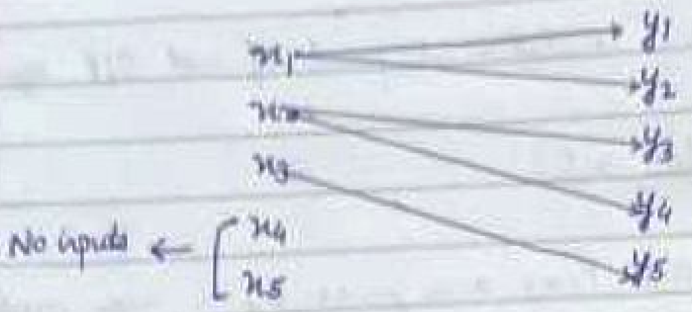
It is called joint probability matrix.

and $P(x_i, y_j)$ is called joint probability of sending x_i and recieving y_j.

- Special chanel

A) A loseless chanel

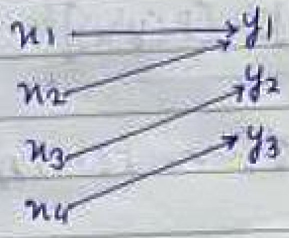
$$[P(Y|X)] = \begin{array}{c|c} & \begin{array}{ccccc} (y_1|x_1) & (y_2|x_1) & (y_3|x_1) & (y_4|x_1) & (y_5|x_1) \end{array} \\ \begin{array}{c} \\ (y_1|x_2) \\ (y_1|x_3) \end{array} & \begin{bmatrix} 1/6 & 5/6 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 3/4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$



- If channel of matrix of a channel contain only one non-zero element in the column, then that channel is defined as lossless channel.

B) Determinastic channel

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



- For determinastic channel, the channel matrix contains only 1 non zero elements in each row. Since each row has one non-zero element this element must be unity.

When a particular input symbol is transmitted through this channel it is known which o/p symbol is recieved

C) Noiseless Chanel

If a chanel is noisy, if it is both determinastic and lossy

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

$x_1 \longrightarrow y_1$

$x_2 \longrightarrow y_2$

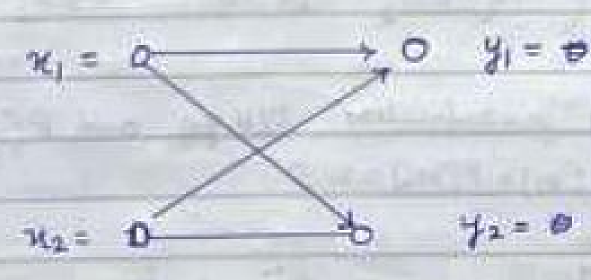
$x_3 \longrightarrow y_3$

$x_4 \longrightarrow y_4$

D) Binary symmetric chanel (BSC)

$$[P(Y/X)] = \begin{bmatrix} 1-P & P \\ P & 1-P \end{bmatrix}$$

It has two inputs and two outputs ; Chanel is symmetric as probability of misinterpreting a transmitted as 0 as 1 is same as misinterpreting 1 as 0

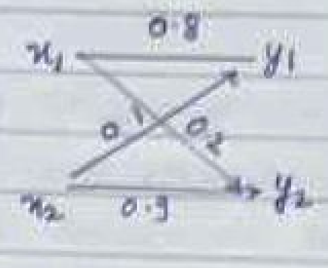


$x_1 = 0$, $y_1 = 0$

$x_2 = 1$, $y_2 = 1$

• $P(x_1) = P(x_2) = 0.5$

Q for the binary channel shown a) find channel matrix
b) $P(y_1), P(y_2)$ when $P(x_1) = P(x_2) = 0.5$



→ $P[Y|X] = \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix}$

$[P(X)] = [0.5 \quad 0.5]$

$[P(Y)] = [0.5 \quad 0.5] \begin{bmatrix} (Y|X) \end{bmatrix}$

$[P(Y)] = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix}$

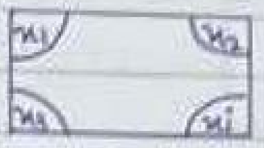
$[P(Y)] = [0.45 \quad 0.55]$

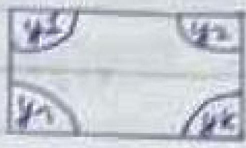
Find the joint probabilities $P(x_1, y_2)$ and $P(x_2, y_1)$
when $P(x_1) = P(x_2) = 0.5$

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

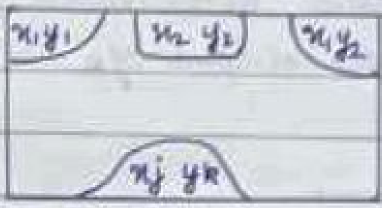
$\begin{bmatrix} P(x_1,y_1) & P(x_1,y_2) \\ P(x_2,y_1) & P(x_2,y_2) \end{bmatrix} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix} = \begin{bmatrix} 0.40 & 0.10 \\ 0.05 & 0.45 \end{bmatrix}$

- Joint entropy and conditional entropy





$S_1 = x_1, x_2, \cdots x_m$ $\qquad$ $S_2 = y_1, y_2 \cdots y_n$



→ Rough representation

$P(X) = P(x_j)$

$P(Y) = P(y_k)$

$P(X,Y) = [P(x_j, y_k)]$

$$\boxed{\because P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B/A_i) P(A_i)}$$

$$H(x) = -\sum_{j=1}^{m} P(x_j) \log_2 P(x_j)$$

$$H(y) = -\sum^{n} P(y_k) \log_2 P(y_k)$$

$$H(x,Y) = -\sum_{j=1}^{m} \sum_{k=1}^{n} P(x_j, y_k) \log_2 P(x_j, y_k)$$

- $H(x)$ and $H(y)$ are called marginal entropy of x and y respectively. The entropy of input symbols x, before their transmission is called as "priori" entropy

- $H(X,Y)$ is joint entropy of (X,Y)

Based on conditional probabilities we can associate entropies

$$\boxed{H(X, y_k) = -\sum_{j=1}^{n} P(x_j/y_k) \log_2 P(x_j/y_k)}$$

The entropy of the input symbol after transmission and reception of particular output y_k is called "posteriori" entropy

$$\boxed{H(X, Y) = \sum_{k=1}^{n} P(y_k)\, H(X/y_k)}$$

Equivocation:-
It is defined as average value of all conditional probability in above equation where $k = 1$ to n

$$\boxed{\text{formula} = -\sum_{k=1}^{n} P(y_k) = \sum_{j=1}^{m} P(x_j/y_k) \log_2 P(x_j/y_k)}$$

$$= -\sum_{j=1}^{m}\sum_{k=1}^{n} P(y_k) \cdot P(x_j/y_k) \log_2 P(x_j/y_k)$$

- $$\boxed{H(X/Y) = -\sum_{j=1}^{m}\sum_{k=1}^{n} P(x_j, y_k) \log_2 P(x_j/y_k)} \quad \ldots ①$$

- $$\boxed{H(Y/X) = -\sum_{j=1}^{m}\sum_{k=1}^{n} \underbrace{P(x_j, y_k)}_{\text{change}} \log_2 P(y_k/x_j)} \quad \ldots ②$$

Equivocation of $(X|Y)$

① It is the measure of $H(X|Y)$. It is the equivocation of x w.r.t y. It is the measure of uncertainity about the channel input. It describes how well one can recover the transmitted symbols from recieved symbols.

② Equivocation of y w.r.t x
Avg uncertainity of channel o/p when x was transmitted and it describes how well one can recover the recieved symbols from transmitted symbols. It is measure of error or noise

Mutual Information

$P(x_j) \rightarrow$ It is the prob that x_j would be selected for transmission. This is priori probability.
Uncertainty in this will be $-\log_2 P(x_j)$. Once y_k is recieved at the o/p, the state of knowledge about x_j is cond^n probability. $P(x_j|y_k)$ which is called posterioric prob.
Uncertainity in this $\rightarrow -\log_2 P(x_j|y_k)$

- $I(x_j, y_k)$ = initial uncertainty - final certainty

- Net reduction in uncertainty

$$I(x_j, y_k) = -\log_2 P(x_j) - [-\log_2 P(x_j|y_k)]$$

$$-\log(x_j) + \log_2 P(x_j|y_k)$$

$$= \log_2 \left(\frac{P(x_j|y_k)}{P(x_j)}\right)$$

$$= \log_2 \left[\frac{P(x_j|y_k)}{P(x_j)} \right]$$

$$= \log_2 \left[\frac{P(x_j, y_k)}{P(x_j)P(y_k)} \right]$$

Q. A binary channel has following properties $[P(Y|X)] = \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix}$

$P(x_1) = 1/4$, $P(x_2) = 3/4$

Find $H(X), H(Y), H(X,Y), H(Y/X)$

$$(P[Y]) = [P(x)] \cdot [P(Y|X)]$$

$$[P(y_1) \cdot P(y_2)] = \left[\frac{1}{4} \;\; \frac{3}{4} \right] \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix}$$

$$= [7/12 \quad 5/12]$$

$$H(X) = -\sum_{j=1}^{2} P(x_j) \log_2 P(x_j) = \frac{1}{4}\log_2 4 + \frac{3}{4}\log_2 \frac{4}{3}$$

$$= \frac{2}{4} + \frac{3}{4} \times 0.41 = 0.811 \text{ bits/msg}$$

$$H(Y) = -\sum_{j=1}^{2} P(y_k) \log_2 P(y_k) = \frac{7}{12}\log \frac{12}{7} + \frac{5}{12}\log_2 \frac{12}{5}$$

$$= 0.976 \text{ bits/msg}$$

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

$$= \begin{bmatrix} 1/4 & 0 \\ 0 & 3/4 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix}$$

$$= \begin{bmatrix} 1/12 & 2/12 \\ 6/12 & 3/12 \end{bmatrix}$$

$H(X,Y) = \sum\sum P(X,Y) \log_2 P(X,Y)$

$$= \frac{1}{12} \log_2 12 + \frac{6}{12} \log_2 \frac{12}{6} + \frac{2}{12} \log_2 \left(\frac{12}{2}\right)$$

$$+ \frac{3}{12} \log_2 \left(\frac{12}{3}\right)$$

$= 1.7299$ bits/msg

Q In a comm^n system, a transmitter has 3 i/p symbols

$A = \{a_1, a_2, a_3\}$
$B = \{b_1, b_2, b_3\}$

The matrix below given shows JPM with some probabilities

$$[P(A,B)] = \begin{bmatrix} 1/12 & * & 5/36 \\ 5/36 & 1/9 & 5/36 \\ * & 1/6 & * \end{bmatrix}$$

$$[P(B)] = \begin{bmatrix} 1/3 & 14/36 & * \end{bmatrix}$$

a) Find missing probabilities (*) in the table

b) Find $P(b_3|a_1), P(a_1|b_3)$

c) Are the events a_1 and b_1 statistically independent? why?

→ Property of JPM:-

1) By adding all elements of JPM column wise we can obtain probability of o/p symbols

$$\sum_{i=1}^{q} P(a_i, b_j) = P(b_j)$$

2) By adding all elements of JPM row-wise we can obtain probability of i/p symbols

$$\sum_{j=1}^{q} P(a_i, b_j) = P(a_i)$$

3) The sum of all elements of JPM is equal to 1

$$\begin{array}{c|ccc} & b_1 & b_2 & b_3 \\ \hline a_1 & 1/12 & 1/9 & 5/36 \\ a_2 & 5/36 & 1/9 & 5/36 \\ a_3 & 1/9 & 1/6 & 0 \end{array}$$

$[P(B)]$ 1/3 14/36 *

→ a) $P(b_1) + P(b_2) + P(b_3) = 1$

$$\frac{1}{3} + \frac{14}{36} + x = 1$$

$$x = 1 - \frac{1}{3} - \frac{14}{36} = \frac{5}{18}$$

Using JPM property 1

$$P(a_1, b_1) + P(a_2, b_1) + P(a_3, b_1) = P(b_1)$$

$$\frac{1}{12} + \frac{5}{36} + x = \frac{1}{3}$$

$$x = \frac{1}{9}$$

$$P(a_1, b_2) + P(a_2, b_2) + P(a_3, b_2) = P(b_2)$$

$$P(a_1, b_2) = \frac{14}{36} - \frac{1}{9} - \frac{1}{6}$$

$$= \frac{1}{9}$$

$$P(a_3, b_3) = 0$$

$$P(a_1, b_1) + P(a_1, b_2) + P(a_1, b_3) = P(a_1)$$

$$P(a_1) = \frac{1}{3}$$

$$P(a_2) = \frac{7}{18}$$

$$P(a_3) = \frac{5}{18}$$

b) $P(b_3 | a_1) = \dfrac{P(a_1, b_3)}{P(a_1)} = \dfrac{5}{12}$

$$P(a_1 | b_3) = \frac{P(a_1, b_3)}{P(b_3)} = \frac{5}{8}$$

c) $P(a_1) \cdot P(b_1) = P(a_1, b_1)$

$$\frac{1}{3} \times \frac{1}{3} \neq \frac{1}{12}$$

It is not statistically independent

V. Imp :-

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

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

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

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

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

- For Lossless channel :-

  chanel capacity $c = \log_2 M \leftarrow$ i/p symbol

- For deterministic

  $c = \log_2 n \leftarrow$ o/p symbols

- For Noisless :-

  $c = \log_2 m = \log_2 n$

- For BSC

  $c = 1 + p\log_2 P + (1-P)\log_2(1-P)$