Information & Communication

ASSIGNMENT - 3

Q1) Given, $F(x) \longrightarrow$ valid CDF

Properties for a function to be a valid CDF are:

i) $F(x)$ should be monotonically non-decreasing.

ii) $F(x)$ should be non-negetive function.

iii) $F(x)$ is right continuous.

iv) $F(-\infty) = 0$ and $F(\infty) = 1$

(a) Let $g(x) = \alpha F(x)$

Try and verify all the above given properties.

i) A function is non decreasing when $\frac{d}{dx} g(x) \geq 0$

$$\frac{d}{dx}(g(x)) = \frac{d}{dx}(\alpha F(x)) = \alpha \frac{d}{dx}(F(x)) > 0$$

[since $F(x)$ is already CDF and non-decreasing.]

$\therefore$ It is monotonically non-decreasing.

(ii) For $g(x)$ to be non-negetive, α must be positive but, $\alpha \in \mathbb{R}$.

Hence, this property isn't satisfied.

(iii) To prove if its right-continuous

$$\lim_{h\to 0} g(x+h) = \lim_{h\to 0} \alpha F(x+h) = \alpha \lim_{h\to 0} F(x+h)$$

$$= \alpha F(x) = g(x)$$

Hence, it is right continuous

(iv) $g(-\infty) = \alpha F(-\infty) = \alpha(0) = 0 \quad [\because F(-\infty) = 0]$

$g(\infty) = \alpha F(\infty) = \alpha(1) = \alpha \quad [\because F(+\infty) = 1]$

$g(\infty) \neq 1$

Hence, it doesn't satisfy this $cond^n$.

$\therefore$ $g(x) = \alpha F(x)$ violates property (ii) and (iv).

$\therefore$ It is not a valid CDF.

(b) $g(x) = F(x)^2$

(i) Check if $\frac{d}{dx}(g(x)) \geq 0$ or not.

$$\frac{d}{dx} g(x) = \frac{d}{dx} (F(x))^2$$

$$= 2F(x) \frac{d}{dx} F(x)$$

Since $F(x)$ is a CDF

$F(x) \geq 0$ (non-negetive)

$\frac{d}{dx} F(x) \geq 0$ (non-decreasing function)

$\Rightarrow \quad \frac{d}{dx}(g(x)) \geq 0$

$\therefore$ It is monotonically non-decreasin function.

(ii) Since $F(x)^2$ is always greater than 0, it is always positive.

$\therefore$ It is non-negetive function.

(iii) $\lim_{h \to 0^+} g(x+h) = \lim_{h \to 0^+} F(x+h)^2$

$= F(x+0)^2$

$= F(x)^2$

$\therefore$ It is right-continuous

(iv) $g(-\infty) = (F(-\infty))^2 = 0^2 = 0$

$g(\infty) = (F(\infty))^2 = 1^2 = 1$

$\therefore$ The limits are satisfied.

Since all properties are satisfied,

$\therefore$ $F(x)^2$ is a valid CDF.

(c) $g(x) = F(x) + (1 - F(x)) \log (1 - F(x))$

(i) Since $F(x)$ is a CDF $\frac{d}{dx} F(x) \geq 0$

Now,

$$\frac{d}{dx}(g(x)) = \cancel{\frac{d}{dx} F(x)} - \frac{d}{dx} F(x) \log(1 - F(x)) + \cancel{\frac{1-F(x)}{1-F(x)} \cdot -\frac{d}{dx} F(x)}$$

$$= -\log(1 - F(x)) \frac{d}{dx} F(x)$$

Wkt,
$$0 \leq F(x) \leq 1$$
$$0 \geq -F(x) \geq -1$$
$$1 \geq 1 - F(x) \geq 0$$
$$\Rightarrow \log(1 - F(x)) \leq 0.$$
$$\Rightarrow -\log(1 - F(x)) \geq 0$$
$$\frac{d}{dx} F(x) \geq 0$$
$$\Rightarrow -\log(1 - F(x)) \frac{d}{dx}(F(x)) \geq 0$$
$$\frac{d}{dx} g(x) \geq 0$$

$\therefore$ It is monotonically non-decreasing function

(ii) Since $F(x)$ is CDF, $F(x) \geq 0$

$$0 \leq F(x) \leq 1 \quad \text{—— } (1)$$
$$0 \geq -F(x) \geq -1$$

$$1+0 \geq 1-F(x) \geq -1+1$$

$$1 \geq (1-F(x)) \geq 0 \quad \text{—— (2)}$$

also, $-\infty \leq \log(1-F(x)) \leq 0$

$$\infty \geq -\log(1-F(x)) \geq 0 \quad \text{—— (3)}$$

2 x 3

$$\infty \geq -(1-F(x))\log(1-F(x)) \geq 0 \quad \text{—— (4)}$$

1 + (4)

$$0 \leq F(x) - (1-F(x))\log(1-F(x)) \leq \infty$$

$$0 \leq g(x) \leq \infty$$

$\therefore$ It is a non-negetive function

(iii) $$\lim_{h\to 0} g(x) = \lim_{h\to 0} F(x+h) + (1-F(x+h))\log(1-F(x+h))$$

$$= \lim_{h\to 0} F(x+h) + \lim_{h\to 0} (1-F(x+h))\log(1-F(x+h))$$

Apply limits

$$= F(x) + (1-F(x))\log(1-F(x))$$

$$= g(x)$$

$\therefore$ It is a right continuous function

(iv) $$g(-\infty) = F(-\infty) + (1-F(-\infty))\log[1-F(-\infty)]$$

as $F(-\infty) = 0$

$$= 0 + (1-0)\log(1-0)$$

$= 0 + (1)(0)$

$g(-\infty) = 0$

$g(\infty) = F(\infty) + (1 - F(\infty)) \log(1 - F(\infty))$

$= 1 + (1 - 1) \log 0$

$= 1 + 0$

$= 1$

$\therefore$ It is satisfied.

Since, all four properties are satisfied.

$\therefore$ $g(x) = F(x) + (1 - F(x)) \log(1 - F(x))$ is a valid CDF.

Q2) Given, X is a continuous random variable.

$$E(x) = \int_{-\infty}^{\infty} x f(x)\, dx$$

(a) $E[X + a]$

$$E[X + a] = \int_{-\infty}^{\infty} (X + a) f(x) dx$$

$$= \int_{-\infty}^{\infty} x f(x)\, dx + a \underbrace{\int_{-\infty}^{\infty} f(x) dx}_{=1 \text{ (cond}^{n} \text{ to be valid pdf)}}$$

$$= E[X] + a$$

Intuitively, we are increasing each and every value of x by a i.e random variable is being shifted by a.

Hence, evenly even the Expected value is increased by a.

(b) $E[ax]$

$$E[ax] = \int_{-\infty}^{\infty} ax\, f(x)\, dx$$

$$= a \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$E[ax] = a\, E[x]$$

Intuitively,

On multiplying the random variable by a, we are changing the scale of measurement by a times. Hence, even the expected value is scaled by a.

(c) $\mathrm{var}[x+a]$

$$\mathrm{var}[x] = E\left(x - E(x)\right)^2$$

$$\mathrm{var}[x+a] = E\left([x+a] - E[x+a]\right)^2$$

$$= E\left(x + \cancel{a} - E(x) - \cancel{a}\right)^2 \quad [\text{from (a)}]$$

$$= E\left(x - E(x)\right)^2$$

$$\mathrm{var}[x+a] = \mathrm{var}[x]$$

$var(X+a) = Var(X)$

Intuitively,

Variation is generally the mean distribution of the random variables across mean. Now adding a constant doesn't change the spread of distribution (evenly cancels out). Hence, the variance remains same.

(d) $var[aX]$

$$var[aX] = E(aX - E[aX])^2$$

$$= E(aX - aE(X))^2$$

$$= E(a^2(X - E(X))^2)$$

$$= a^2 E(X - E(X))^2 \quad [\text{from (b)}$$

$$var[aX] = a^2 var(X)$$

Intuitively,

When X is multiplied with a, the mean gets scaled by a. And in variance, we have a term i.e, the square of mean. Hence even the variance gets scaled by a^2.

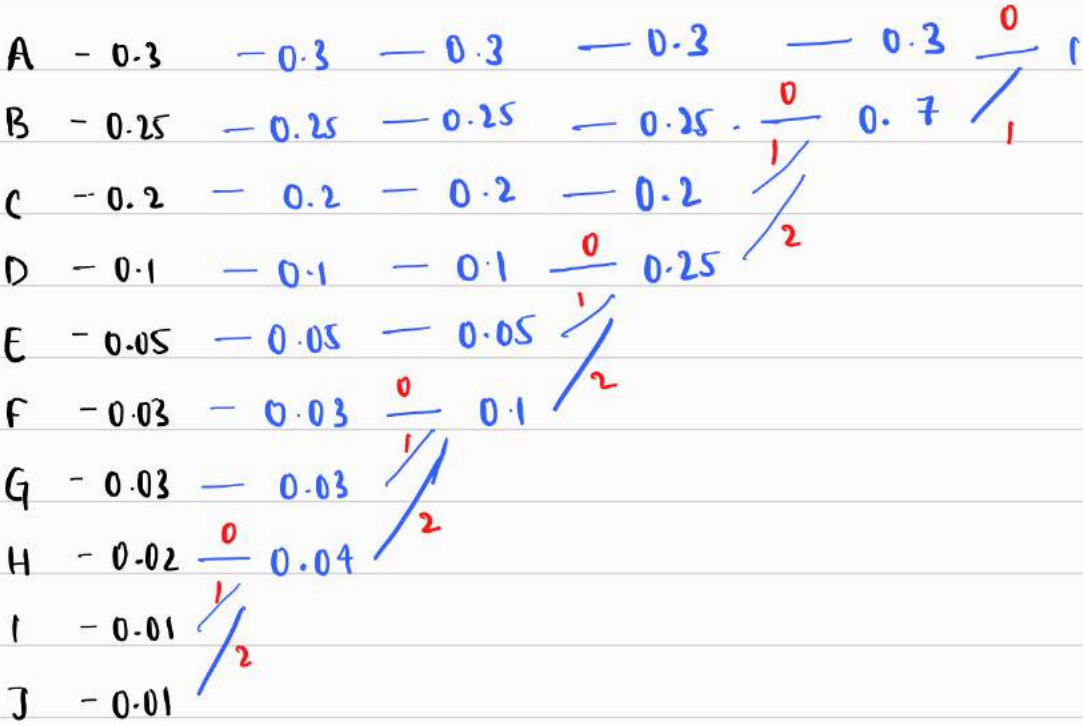
Q8) Given ... $S = \{A, B, C, D, E, F, G, H, I, J\}$

Given source S = {A, B, C, D, E, F, G, H, I, J}

The probabilities of all the symbols are given:

P(A) = 0.3	P(F) = 0.03
P(B) = 0.25	P(G) = 0.03
P(C) = 0.2	P(H) = 0.02
P(D) = 0.1	P(I) = 0.01
P(E) = 0.05	P(J) = 0.01

(a) Ternary Huffmann Coding



Source Code Table:

Symbol	Codeword	Length
A	0	1
B	10	2
C	11	2
D	120	3
E	121	3
F	1220	4

G	1221	4
H	12220	5
I	12221	5
J	12222	5

(b) Expected Length $L(C)$:

$$L(C) = \sum_{x \in \mathcal{X}} p(x)\, L(x)$$

$$= 0.3(1) + (0.25)(2) + (0.2)(2) + (0.1)(3) + 0.05(3) + 0.03(4) + 0.03(4) + (0.02)(5) + (0.01)(5) + (0.01)(5)$$

$$= 0.3 + 0.5 + 0.4 + 0.3 + 0.15 + 0.12 + 0.12 + 0.1 + 0.05 + 0.05$$

$$\boxed{L(C) = 2.09}$$

(c) Efficiency

$$\eta = \frac{H(X)}{L(C)}$$

$$H(X) = -\sum_{x \in X} p(x) \log_3 (p(x))$$

Note: Base = 3

$$= -\Big[0.3 \log_3(0.3) + 0.25 \log_3(0.25) + 0.2 \log_3(0.2) + 0.1 \log_3(0.1) + 0.05 \log_3(0.05) + 0.03 \log_3(0.03) + 0.03 \log_3(0.03) + 0.02 \log_3(0.02) + 0.01 \log_3(0.01) + 0.01 \log_3(0.01)\Big]$$

$$= -(-1.69162706) = 1.69162706$$

$$\eta = \frac{H(X)}{\ldots} = \frac{1.6916}{\ldots} \Rightarrow \boxed{\eta = 0.8093}$$

$L(C)$ 2.09 $\Rightarrow$

Q9) Given

Source X with $x_i : i \in \{1, 2, 3, 4\}$

$P(1) = 0.4$ $P(3) = 0.2$

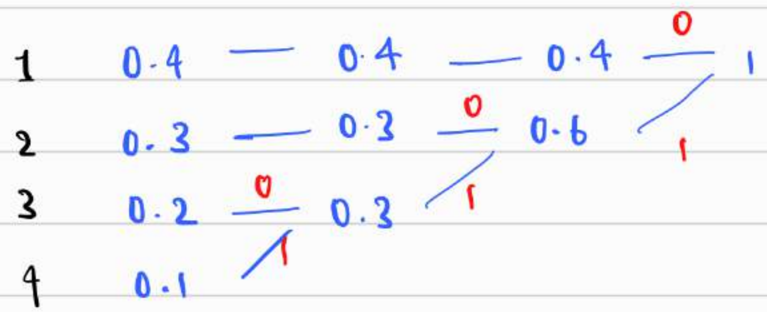
$P(2) = 0.3$ $P(4) = 0.1$

(a) Entropy $= -\sum_{x \in X} p(x) \log p(x)$

$$= -(0.4)\log(0.4) - (0.3)\log(0.3) - (0.2)\log(0.2) - (0.1)\log(0.1)$$

$$= +1.8461$$

(b) Huffmann Coding



Symbol	Source Code	Length
1	0	1
2	10	2
3	110	3
4	111	3

Expected Length, $L(C) = \sum p(x) L(x)$

$x \in \mathcal{X}$

$= (0.4)(1) + (0.3)(2) + (0.2)(3) + (0.1)3$

$= 0.4 + 0.6 + 0.6 + 0.3$

$$\boxed{L_1 = 1.9}$$

(c) Given symbols x_1, x_2

$$P_{x_1, x_2} = p(x_1) \cdot p(x_2)$$

Joint Distribution of P_{x_1, x_2}

$x_1 \backslash x_2$	1	2	3	4
1	0.16	0.12	0.08	0.04
2	0.12	0.09	0.06	0.03
3	0.08	0.06	0.04	0.02
4	0.04	0.03	0.02	0.01

Huffmann Coding

0.01 0.02 0.02 0.03 0.03 0.04 0.04 0.04 0.06 0.06 0.08 0.08 0.09 0.12 0.12 0.16

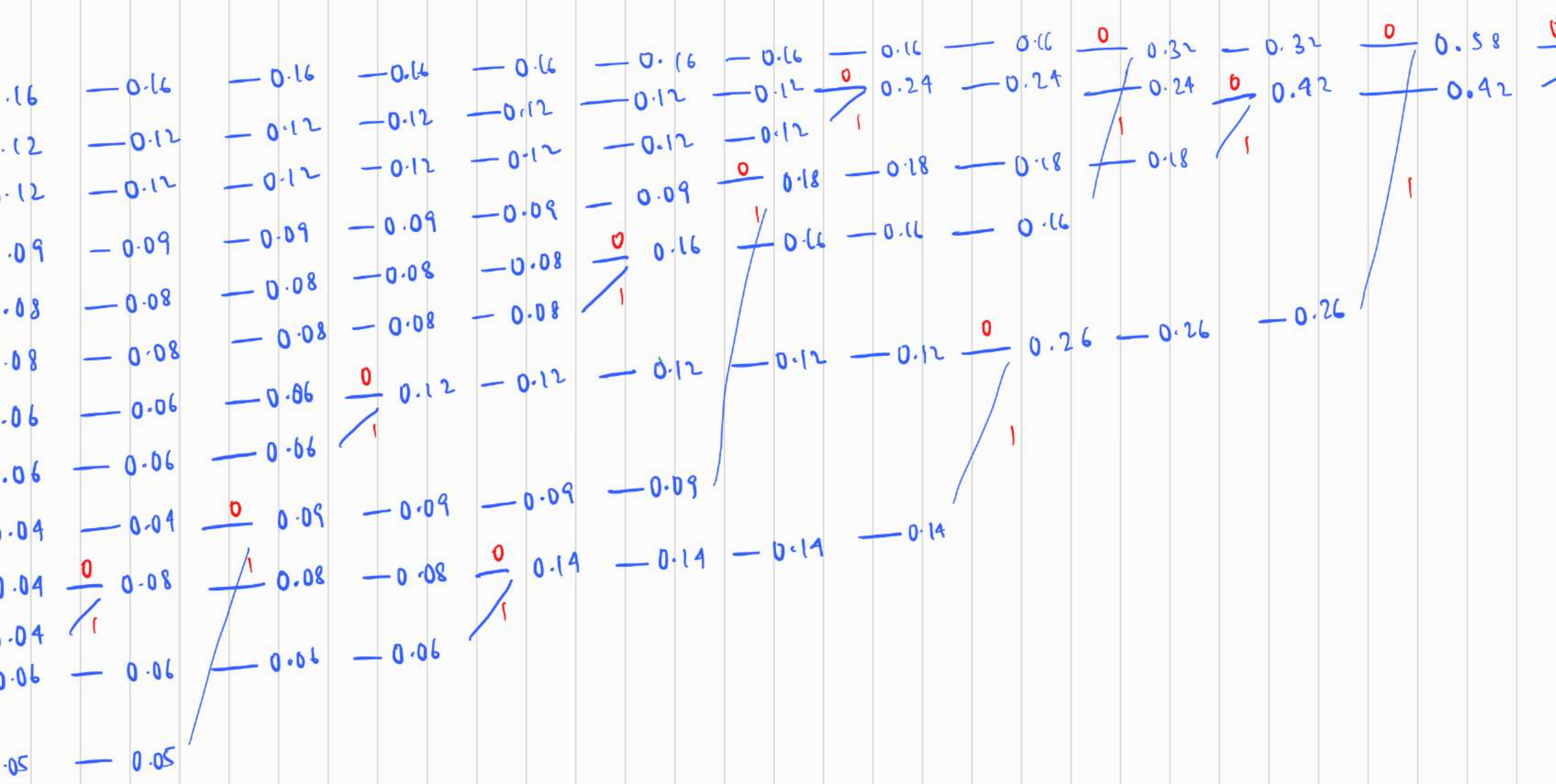
1 0

0.03 0.02 0.03 0.03 0.04 0.04 0.04 0.06 0.06 0.08 0.08 0.09 0.12 0.12 0.16

1 0

0.05 0.03 0.03 0.04 0.04 0.04 0.06 0.06 0.08 0.08 0.09 0.12 0.12 0.16

1 0



Symbol	P	CodeWord	Length
p(1,1)	0.16	000	3
p(1,2)	0.12	100	3
p(1,3)	0.08	0010	4
p(1,4)	0.04	01100	5
p(2,1)	0.12	101	3
p(2,2)	0.09	110	3
p(2,3)	0.06	0100	4
p(2,4)	0.03	01110	5
p(3,1)	0.08	0011	4
p(3,2)	0.06	0101	4
p(3,3)	0.04	1110	4
p(3,4)	0.02	11110	5
p(4,1)	0.04	01101	5
p(4,2)	0.03	01111	5
p(4,3)	0.02	111110	6
p(4,4)	0.01	111111	6

Expected Length, $L(C) = \sum_{x \in \mathcal{X}} p(x) . L(x)$

$$L(C) = (0.16)(3) + (0.12)(3) + (0.08)(4) + (0.04)5 + (0.12)(3) + (0.09)(3) + (0.06)(4) + (0.03)5 + (0.08)4 + (0.06)(4) + (0.04)(4) + (0.02)5 + (0.04)5 + (0.03)5 + (0.02)6 + (0.01)6$$

$$\boxed{L_2 = 3.73}$$

(e) **Comparision :**

$$L_1' = \frac{L_1}{1} = 1.9$$

$$L_2' = \frac{L_2}{2} = \frac{3.73}{2} = 1.865$$

$$H(X) = 1.8461$$

$$\eta_1 = \frac{H(X)}{L_1'} = \frac{1.8461}{1.9} = 0.9716$$

$$\eta_2 = \frac{H(X)}{L_2'} = \frac{1.8461}{1.865} = 0.9898$$

Here,

$$L_1 > L_2 > H(X)$$

From the above given we can say that,

For higher order extensions, the expected length of the source code decreases.

Because of this the coding efficiency increases ↑

$$\eta_1 < \eta_2 .$$

This property of extensions can be utilised to obtain $L(C) = H(X)$.
i.e, keep increasing the no. of extentions until the value of $L(C)$ gets down to $H(X)$.
Then we can obtain maximum coding efficiency.

(f) When 3-iid symbols are compressed using Huffmann Coding, then the expected length of the source code will decrease when compared to L_1 and L_2

i.e, $L_3 < L_2 < L_1$

This concludes that, as extension increases, the expected length decreases.
Also, L_3 will be closer to $H(X)$ but little greater

$\Rightarrow$ $H(X) < L_3 < L_2 < L_1$

Q3) Given, 7-match cricket series
Two teams A and B

$X \rightarrow$ outcomes of series after completion
$Y \rightarrow$ no. of games played (range from 4 to 7)

(a) $H(Y|X)$

$H(Y|X)$ – measure the uncertainity in the no. of games that are played given the information about outcome of series (X).

When we already know the outcome of the series, then we will have no uncertainity in knowing the no. of games played as it is already evident in X. Hence uncertainity in Y given X is zero.
Hence, intuitively $H(Y|X)$ must be 0.

(b) $H(X)$ - measures the uncertainity of outcome of the series. It represents all possible outcomes of the series before any game.

$H(Y)$ - measures the uncertainity in the no. of games played. It tells us about the number of games that are going to be played.

→ X deals with outcome of series, which contains various different wins and losses of teams A once, and B once and this leads to higher uncertainity. Because of this the no. of possibilities increases exponentially as Y increases.

Y is just ranging from games 4 to 7 which is a lot constrained when compared to the uncertainity in X.
Hence, entropy of X is greater than entropy of Y.

$$H(X) > H(Y)$$

(c) All the possible outcomes of the series are listed below:

When $Y = 4 \longrightarrow$ AAAA, BBBB $(1 + 1 = 2$ ways$)$

$Y = 5 \longrightarrow$ BAAAA ABBBB $(4 + 4 = 8$ ways$)$

ABAAA BABBB

AABAA BBABB

AAABA BBBAB

similarly

$Y = 6 \longrightarrow 10 + 10 = 20$ ways

$Y = 7 \longrightarrow 20 + 20 = 40$ ways

Now, consider only A is winning the series,

if $Y =$	4	5	6	7
No. of dif possibilities	1	${}^4C_1 = 4$	${}^5C_2 = 10$	${}^6C_3 = 20$
Probability of each case	$= \frac{1}{2^4}$	$= \frac{1}{2^5}$	$= \frac{1}{2^6}$	$= \frac{1}{2^7}$

This is the same for the case when B is winning.

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

PMF of X $= \left[\left(\frac{1}{2^4}\right) \times 2 \text{ times}, \left(\frac{1}{2^5}\right) \times 8 \text{ times}, \left(\frac{1}{2^6}\right) \times 20 \text{ times}, \left(\frac{1}{2^7}\right) \times 40 \text{ times} \right]$

$$H(X) = -\sum_{x \in \mathcal{X}} p_x(X=x) \log_2 P_X(X=x)$$

$$= -\left[2 \times \frac{1}{2^4} \log_2\left(\frac{1}{2^4}\right) + 8 \times \frac{1}{2^5} \log_2\left(\frac{1}{2^5}\right) + 20 \times \frac{1}{2^6} \log_2(1/2^6) + 40 \frac{1}{2^7} \log_2(1/2^7)\right]$$

$$= \left(2 \times \frac{1}{2^4} \times 4\right) + \left(8 \times \frac{1}{2^5} \times 5\right) + \left(20 \times \frac{1}{2^6} \cdot 6\right) + \left(40 \frac{1}{2^7} \times 7\right)$$

$$= 0.5 + 1.25 + 1.875 + 2.1875$$

$$\boxed{H(X) = 5.8125}$$

Total no. of possible cricket series $= 2(1 + 4 + 10 + 20)$
$= 70$

PMF of y $= \left[\frac{2}{70}, \frac{8}{70}, \frac{20}{70}, \frac{40}{70}\right]$

$$= \left[\frac{1}{35}, \frac{4}{35}, \frac{10}{35}, \frac{20}{35}\right]$$

$$H(y) = -\sum_{y \in \mathcal{Y}} P_Y(Y=y) \log_2 P_Y(Y=y)$$

$$= -\left[\frac{1}{35} \log_2(1/35) + \frac{4}{35} \log_2(4/35) + \frac{10}{35} \log_2(10/35) + \frac{20}{35} \log_2(20/35)\right]$$

$$= \frac{1}{35}\left[35 \log_2 35 - 4 \log_2 4 - 10 \log_2 10 - 20 \log_2 20\right]$$

$$\boxed{H(Y) = 1.48 \text{ bits}}$$

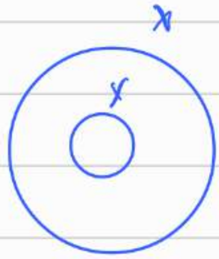
$$H(X|Y) = -\sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p(y,x) \log\left[p(x|y)\right]$$

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

$$= H(X) - H(Y)$$

$$= 5.8125 - 1.48$$

$$\boxed{H(X|Y) = 4.3325 \text{ bits}}$$



In $H(X,Y)$, as we already have X, no need of Y

$$H(Y|X) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log\left[p(y|x)\right]$$

$p(y=4 / x=x) = 1 \quad \exists \; x : x = AAA, BBB$

$p(y=5 / x=x) = 1 \quad \exists \; x : AABA, \ldots ABBBB$

$p(y=6 / x=x) = 1 \quad \exists \; x : AABBA, \ldots AABBBB$

$p(y=7 / x=x) = 1 \quad \exists \; x : AAABBBA, \ldots AAABBB$

These when combined with ①

$$H(Y|X) = 0 \text{ bits}$$

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

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

$$\boxed{I(X;Y) = 1.48 \text{ bits}}$$

Q4) To prove :

$$2\,H(X,Y,Z) \le H(X,Y) + H(Y,Z) + H(X,Z)$$

$$2\,H(X,Y,Z) = 2\,H(X) + 2\,H(Y|X) + 2\,H(Z|Y,X)$$

(by chain rule of entropy) — (1)

We know that,

"Coditioning does not increase entropy"

$\Rightarrow \quad H(Y|X) \le H(Y)$

$$H(Z|X,Y) \le H(Z|X) \le H(Z) \quad \text{— (3)}$$

$$H(Z|X,Y) \le H(Z|Y) \le H(Z) \quad \text{— (4)}$$

$$H(X,Y) = H(X) + H(Y|X) \quad \text{— (2)}$$

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

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

Try rearranging eq (1).

$$2\,H(X,Y,Z) = \big(H(X) + H(Y|X)\big) + \big(H(X) + H(Z|X,Y)\big) +$$

$$\left(H(Y|X) + H(Z|X,Y)\right)$$

$$H(X) + H(Y|X) = H(X,Y) \quad \text{(5)} \qquad \text{from (2)}$$

$$H(X) + H(Z|X,Y) \le H(X) + H(Z|X) \qquad \text{(from 3)}$$

$$\le H(X,Z) \quad \text{(6)}$$

$$H(Y|X) + H(Z|X,Y) \le H(Y) + H(Z|Y) \qquad \text{(from 4)}$$

$$\le H(Y,Z) \quad \text{(7)}$$

Add 5, 6, 7

$$\Rightarrow H(X) + H(Y|X) + H(X) + H(Z|X,Y) + H(Y|X) + H(Z|X,Y) \le H(X,Y) + H(X,Z) + H(Y,Z)$$

from (1)

$$2\,H(X,Y,Z) \le H(X,Y) + H(X,Z) + H(Y,Z)$$

Q 5) Given,

$X, Y \longrightarrow$ independent binary R.V

$X, Y \longrightarrow \{0,1\}$

$$Z = X, Y \quad \Rightarrow \quad Z \in \{0,1,2\}$$

X, Y are independent $\Rightarrow H(X,Y) = H(X) + H(Y)$

$I(X;Y) \longrightarrow$ gives mutual information b/w X, Y.

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

$$= H(X) - H(X)$$

$$I(X;Y) = 0$$

$\therefore$ No mutual information b/w X and Y.
i.e, each of them can't give any information about the other.

$I(X;Y|Z) \longrightarrow$ mutual information b/w X and Y given the value of Y.

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

$$= -\sum_{x\in X}\sum_{z\in Z} p(x,z)\log_2[p(x|z)] + \sum_{y\in Y}\sum_{z\in Z} p(y,z)\, H(X/Y=y, Z=z)$$

Now, when $p(y,z)=0$, the right term becomes 0
and when,
$p(y,z) \neq 0$, X is already know i.e $X = Z - Y$.
$\Rightarrow$ $H(X/Y=y, Z=z) = 0$

$$I(X;Y/Z) = -\sum_{x\in X}\sum_{z\in X} p(x,z)\log_2[p(x|z)]$$

wkt $p(0,2) = p(1,0) = 0$

$$p(0,0) = p(0,1) = p(1,1) = p(1,2) = 1/4$$

$\therefore$ $p(0|0) = p(1|2) = 1$
$p(0|1) = p(1|1) = 1/2$

$\Rightarrow \quad I(X;Y|Z) = -\left[\frac{1}{4}\log_2 1 + \frac{1}{4}\log_2 1 + \frac{1}{4}\log_2 \frac{1}{2} + \frac{1}{4}\log_2 \frac{1}{2}\right]$

$= 0.5$ bits

Clearly, $\quad I(X;Y/Z) > I(X;Y)$

Now, when $X = Y$

$$I(X;Y) = H(X) - \cancel{H(X/Y)}^{\,0} = 2 \cdot \frac{1}{2}\log_2 2$$

$$= 1 \text{ bit}$$

$$I(X;Y|Z) = H(X|Z) - H(X/Y,Z)$$

since $X = Y$, $Z = X + Y$ can hold values 0, 1, 2 only.

If $Z = 0$, $X = 0$, $Y = 0$; if $Z = 2$, $X = 1$ & $Y = 1$

$\Rightarrow \quad H(X|Z) = 0 = H(X|Y,Z)$

$\Rightarrow \quad I(X;Y,Z) = 0$ bits.

The previous relationship is no longer valid,
Now,

$$I(X;Y) > I(X;Y/Z)$$

Q6)

(a) $H(X_1, X_2, X_3, \ldots, X_n) \leq \sum_{i=1}^{n} H(X_i)$

We know a property that says,

" Conditioning does not increase entropy."

LHS $= H(X_1, X_2, X_3, \ldots, X_n)$

By chain rules,

$$= \sum_{i=1}^{n} H(X_i \mid X_{i-1}, X_{i-2}, \ldots X_1)$$

$$= H(X_1) + H(X_2 \mid X_1) + H(X_3 \mid X_2, X_1) + \ldots + H(X_n \mid X_{n-1}, \ldots X_1)$$

These are all the conditional entropies.

On RHS, we have individual entropies of all X_i.

WKT
$$H(X_2) \geq H(X_2 \mid X_1)$$
$$H(X_3) \geq H(X_3 \mid X_2, X_1)$$
$$\vdots$$
$$H(X_n) \geq H(X_n \mid X_{n-1} \ldots X_1)$$

On adding all. we get

$$\sum_{i=1}^{n} H(X_i) \geq H(X_1, X_2, X_3, \ldots X_n)$$

Equality holds true only when $X_1, X_2, \ldots, X_n$ are independent.

(b) $I(X;Y) \geq 0$

(b) $I(X;Y) \geq 0$

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

$$= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log p(x|y)$$

$$= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log\left(\frac{p(x,y)}{p(x)\,p(y)}\right)$$

Let $p(x,y) = u(a)$

$p(x)\,p(y) = v(a)$

$$I(X;Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} u(a) \log \frac{u(a)}{v(a)}$$

$$= D(u \| v)$$

[From theorem given, $D(p\|q) \geq 0$]

$\therefore \quad I(X;Y) \geq 0$

Equality holds if X & Y are independent RV

i.e. $H(X) = H(X|Y)$

(c) $H(X) \leq \log |\mathcal{X}|$

$|\mathcal{X}| \longrightarrow$ cardinality of set $\mathcal{X}$.

$$D(p\|q) = \sum_{x} p(x) \log \frac{p(x)}{q(x)}$$

let $p \rightarrow$ pmf of X.

$q \rightarrow$ uniform distribution over $\mathcal{X}$.

Let us take,

$$q(x) = \frac{1}{|\mathcal{X}|} \quad \forall\, x \in \mathcal{X}$$

On substituting

$$D(p\|q) = \mathbb{E}_p \log \frac{p(x)}{q(x)}$$

$$= \mathbb{E}_p\left[\log p(x) + \log|\mathcal{X}|\right]$$

$$= \mathbb{E}_p\left(\log|\mathcal{X}|\right) - \left(-\mathbb{E}_p \log p(x)\right) \geq 0$$

$$\left[\because D(p\|q) \geq 0\right]$$

$$\Rightarrow \quad H(x) \leq \log|\mathcal{X}|$$

the equality holds iff $p(x) = q(x)$ i.e, X assumes uniform distribution over all its realisations $\frac{1}{|\mathcal{X}|}$.

Q7)

(a) $H(X, Y, Z \mid W_1, W_2, W_3)$

$$= H(X \mid W_1, W_2, W_3) + H(Y \mid W_1, W_2, W_3, X) +$$

$$H(Z \mid W_1, W_2, W_3, X, Y)$$

(b) $I(X_1, X_2 ; Z_1 Z_2 \mid Y_1 Y_2)$

$$= H(X_1 X_2 \mid Y_1, Y_2) - H(X_1 X_2 \mid Y_1, Y_2, Z_1, Z_2)$$

$$= H(X_1 \mid Y_1, Y_2) + H(X_2 \mid Y_1 Y_2 X_1)$$

$$- H(X_1 \mid Y_1, Y_2, Z_1, Z_2) - H(X_2 \mid Y_1, Y_2, Z_1, Z_2, X_1)$$

$H(x_1 | y_1, y_2, z_1, z_2) \quad H(x_2 | y_1, y_2, z_1 z_2 x_1)$

Q8), Q9) are written after Q2).