SHANNON - FANO CODING.

Procedure :-

1. List the source symbols in order of decreasing probability.

2. Partition the set into two sets that are as close to equiprobable as possible, and assign '0' to the upper set and '1' to the lower set.

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

Problem :-

A DMS has five symbols x_1, x_2, x_3, x_4 and x_5 with respectively probabilities 0.4, 0.19, 0.16, 0.15 and 0.15 attached to every symbol.

(i) Construct a Shannon - Fano code for the source and calculate code efficiency.

(ii) Construct a Huffman code for the source and calculate code efficiency.

(iii) Compare the code efficiency of the two techniques.

Step 1 :-

Arranging the probabilities in descending order.

Source	probability
x_1	0.4
x_2	0.19
x_3	0.16
x_4	0.15
x_5	0.15

Step 2 :- Partition the set into two ~~i~~ sets — as close to equiprobable as possible.

and assign '0' to upper set and '1' to lower set.

I

		I	
x_1	0.4	0	Upper set.
	0.4 / 0.65		
x_2	0.19	1	lower set.
x_3	0.16	1	
x_4	0.15	1	
x_5	0.15	1	

Step 3:- Need to Continue the process of partitioning, until further partitioning is not possible.

			I	II	
x_1	0.4		0	X	
x_2	0.19	$\oplus$	1	0	upper set
x_3	0.16		1	0	
		0.35 / 0.3.			
x_4	0.15	$\oplus$	1	1	lower set.
x_5	0.15		1	1	

step 5:- Again iterating.

		I	II	III	Codeword	
x_1	0.4	0	x	x	0	Partition 1
x_2	0.19	1	0	0	100	Partition 3
x_3	0.16	1	0	1	101	Partition 2
x_4	0.15	1	1	0	110	Partition 4.
x_5	0.15	1	1	1	111	

0.19 / 0.16

0.15 / 0.15

Entropy $H(x) = -\sum_{i=1}^{n} P_i \log_2 P_i$.

$= 2.28$ bits/symbol .

Average code length

$$\bar{N} = \sum_{i=1}^{n} P_i n_i$$

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

$$= 2.35$$

Efficiency $\eta = \frac{2.28}{2.35} \times 100\%$

$= 94.8\%$