

Zadatak 4.) Sedam izvorišnih simbola pojavljuje se s vjerojatnostima 1/4, 1/5, 1/6, 1/6, 1/12, 1/12, 1/20, a treba ih kodirati uporabom četiri kodna simbola a, b, c i d. Odredite: (a) kôd primjenom Shannon-Fano metode; (b) efikasnost dobivenog koda.

[R. (b) 0.85 (bolje kodiranje daje 0.95; druga podjela simbola u skupine)]

Treba poredati izvorišne simbole silazno po vjerojatnostima i grupirati ih u 4 skupine sa podjednakim vjerojatnostima. Zatim simbole koji spadaju u iste skupine ponovno dijeliti i tako redom dok ne dođemo do naših kodnih riječi:

Xn	P(Xn)	Korak 1	Korak 2	Kodna riječ	Duljina kodne riječi
1	1/4	a		a	1
2	1/5	b		b	1
3	1/6	c	a	ca	2
4	1/6	c	b	cb	2
5	1/12	d	a	da	2
6	1/12	d	b	db	2
7	1/20	d	c	dc	2

Efikasnost dobivenog koda se računa formulom:

$$\varepsilon = \frac{H(X)}{L * \log_2 r}$$

Gdje je r broj kodnih simbola (najčešće kodiramo u 1 i 0, gdje je $r = 2$ što nam daje $\varepsilon = \frac{H(X)}{L}$).

Mi imamo a,b,c i d, što znači $r = 4$.

Prosječna duljina L se računa preko formule:

$$L = \sum [P(Xn) * (Duljina kodne riječi)].$$

$$L = 1*1/4 + 1*1/5 + 2*1/6 + 2*1/6 + 2*1/12 + 2*1/12 + 2*1/20$$

$$L = 1.55$$

Entropija se računa preko formule:

$$H = \sum [P(Xn) * \log(Xn)]$$

$$H = - \left(\frac{1}{4} \log \frac{1}{4} + \frac{1}{5} \log \frac{1}{5} + \frac{1}{6} \log \frac{1}{6} + \frac{1}{6} \log \frac{1}{6} + \frac{1}{12} \log \frac{1}{12} + \frac{1}{12} \log \frac{1}{12} + \frac{1}{20} \log \frac{1}{20} \right) =$$

$$H = 2.64$$

I sad samo uvrstimo u formulu:

$$\varepsilon = \frac{H(X)}{L * \log_2 r} = \frac{2.64}{1.55 * \log_2 4} = \frac{2.64}{1.55 * 2} = 0.85$$