Άσκηση 3 (Φυλλάδιο 2η)

Μία πηγή πληροφορίας παράγει 8 διαφορετικά σύμβολα (u_1, u_2,…,u_8) με αντίστοιχες πιθανότητες 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128, 1/128. Αυτά τα σύμβολα κωδικοποιούνται σαν 000, 001, 010, 011, 100, 101, 110, 111 αντίστοιχα.

(α) Ποιο είναι το ποσό πληροφορίας για κάθε σύμβολο;

(β) Ποιες είναι οι πιθανότητες εμφάνισης ενός 0 και ενός 1;

(γ) Ποια είναι η αποδοτικότητα αυτού του κώδικα;

(δ) Δώστε έναν αποδοτικό κώδικα με τη βοήθεια της μεθόδου Fano ή του Shannon.

(ε) Ποια είναι η αποδοτικότητα του κώδικα που προκύπτει;

▶

(α) Το ποσό πληροφορίας για κάθε σύμβολο εξαρτάται από την πιθανότητα που αυτό εμφανίζεται στον κώδικα παρά από αυτό το ίδιο το σύμβολο.
Έτσι έχουμε : $H(u_1)$=-(1/2)log(1/2)=0.5, $H(u_2)$=-(1/4)log(1/4)=0.5,
$H(u_3)$=-(1/8)log(1/8)=0.375, $H(u_4)$=-(1/16)log(1/16)=0.25 κλπ.
Τότε το ποσό πληροφορίας της πηγής βρίσκεται από την σχέση :

$$H(U)=-\sum_{i=1}^{8}p_i\log p_i=-\frac{1}{2}\log\frac{1}{2}-\frac{1}{4}\log\frac{1}{4}-\frac{1}{8}\log\frac{1}{8}-\frac{1}{16}\log\frac{1}{16}-\frac{1}{32}\log\frac{1}{32}-$$
$$-\frac{1}{64}\log\frac{1}{64}+\frac{2}{128}\log\frac{1}{128}=1.9844\ \text{bits}/\sigma\upsilon\mu\beta o\lambda o.$$

(β) p(εμφάνισης ενός μόνο 0) = p(011) + p(101) + p(110) =

$$=\frac{1}{16}+\frac{1}{64}+\frac{1}{128}=0{,}0859.$$

p(εμφάνισης ενός μόνο 1) = p(001) + p(010) + p(100) =

$$=\frac{1}{4}+\frac{1}{8}+\frac{1}{32}=0{,}4063.$$

ενώ η πιθανότητα εμφάνισης ενός μόνο 0 και 1 συγχρόνως, είναι 0 στις δοθείσες κ.λ..

(γ) Η αποδοτικότητα ενός κώδικα βρίσκεται από τη σχέση :

$$\eta = \frac{H(U)}{L \cdot \log r},$$ όπου L το μέσο μήκος του κώδικα,

άρα στην περίπτωσή μας : $L = \sum_{i=1}^{8} p_i \ell_i = \sum_{i=1}^{8} 3 \cdot p_i = 3 \cdot \sum_{i=1}^{8} p_i = 3 \cdot 1 = 3$ και

r=2, άρα η αποδοτικότητά του θα είναι : $\eta = \frac{H(U)}{L \cdot \log 2} = \frac{1.9844}{3} = 66.15\%$.

(δ) Μέθοδος Fano:

Τα σύμβολα της πηγής είναι ήδη διατεταγμένα κατά φθίνουσα διάταξη σύμφωνα με τις πιθανότητές τους.

Χωρίζουμε το σύνολο των συμβόλων σε r=2 υποσύνολα κάθε φορά έτσι ώστε το άθροισμα των πιθανοτήτων σε κάθε σύνολο να είναι το ίδιο ή σχεδόν το ίδιο:

$\{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$
$\{u_1\} \to 0$ $\{u_2, u_3, u_4, u_5, u_6, u_7, u_8\} \to 1$
$\{u_2\} \to 0$ $\{u_3, u_4, u_5, u_6, u_7, u_8\} \to 1$
$\{u_3\} \to 0$ $\{u_4, u_5, u_6, u_7, u_8\} \to 1$
$\{u_4\} \to 0$ $\{u_5, u_6, u_7, u_8\} \to 1$
$\{u_5\} \to 0$ $\{u_6, u_7, u_8\} \to 1$
$\{u_6\} \to 0$ $\{u_7, u_8\} \to 1$
$\{u_7\} \to 0$ $\{u_8\} \to 1$

Επομένως οι κωδικές λέξεις κάθε συμβόλου είναι :
u_1 : 0
u_2 : 10
u_3 : 110
u_4 : 1110
u_5 : 11110
u_6 : 111110
u_7 : 1111110
u_8 : 1111111

Το μέσο μήκος του κώδικα είναι :

$$L=\sum_{i=1}^{8}p_i\ell_i=1\cdot\frac{1}{2}+2\cdot\frac{1}{4}+3\cdot\frac{1}{8}+4\cdot\frac{1}{16}+5\cdot\frac{1}{32}+6\cdot\frac{1}{64}+7\cdot\frac{1}{128}+7\cdot\frac{1}{128}=$$

=1,9844

επομένως ο κώδικας αυτός θα έχει αποδοτικότητα ίση με

$\eta=\frac{H(U)}{L\cdot\log 2}=\frac{1.9844}{1.9844}=1$, δηλαδή είναι ένας βέλτιστος κώδικας, γεγονός που θα περίμενε κανείς αν είχε παρατηρήσει ότι τα μήκη των κωδικών λέξεων είναι τέτοια ώστε να ισχύει : $p_i=\left(\frac{1}{2}\right)^{\ell_i}$, i=1,2,...8.

Μέθοδος Shannon:

Σύμφωνα με τη μέθοδο Shannon, υπολογίζουμε τις αθροιστικές πιθανότητες p_i , βρίσκουμε το μήκος κάθε κωδικής λέξης και γράφουμε σε δυαδική μορφή τα p_i , οπότε καταλήγουμε στον παρακάτω πίνακα:

Σύμβολο	$p(u_i)$	p_i	ℓ_i	Κώδικας
u_1	1/2	0	1	0
u_2	1/4	1/2	2	10
u_3	1/8	3/4	3	110
u_4	1/16	7/8	4	1110
u_5	1/32	15/16	5	11110
u_6	1/64	31/32	6	111110
u_7	1/128	63/64	7	1111110
u_8	1/128	127/128	7	1111111

Παρατηρούμε ότι ο κώδικας που προκύπτει με τη μέθοδο Shannon είναι ακριβώς ο ίδιος με αυτόν της μεθόδου Fano.
Προφανώς, η αποδοτικότητα θα είναι και πάλι ίση με τη μονάδα. (η=1)