

Практичне завдання №4

СТАТИСТИЧНІ МЕТОДИ ЕКОНОМНОГО КОДУВАННЯ

Кравець Ольга ПМО-21

Варіант 9

4.1. Побудувала нерівномірні ефективні коди за алгоритмами Шеннона-Фано та Хаффмена. Порівняла ефективність кодів

Варіант	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9
9	0,28	0,35	0,1	0,03	0,06	0,04	0,04	0,04	0,06

Код за алгоритмом Шеннона-Фано:

p_2	0,35	11	2	0,7
p_1	0,28	10	2	0,56
p_3	0,1	011	3	0,3
p_5	0,06	0101	4	0,24
p_9	0,06	0100	4	0,24
p_6	0,04	0011	4	0,16
p_7	0,04	0010	4	0,16
p_8	0,04	0001	4	0,16
p_4	0,03	0000	4	0,12
\tilde{l}				2,64

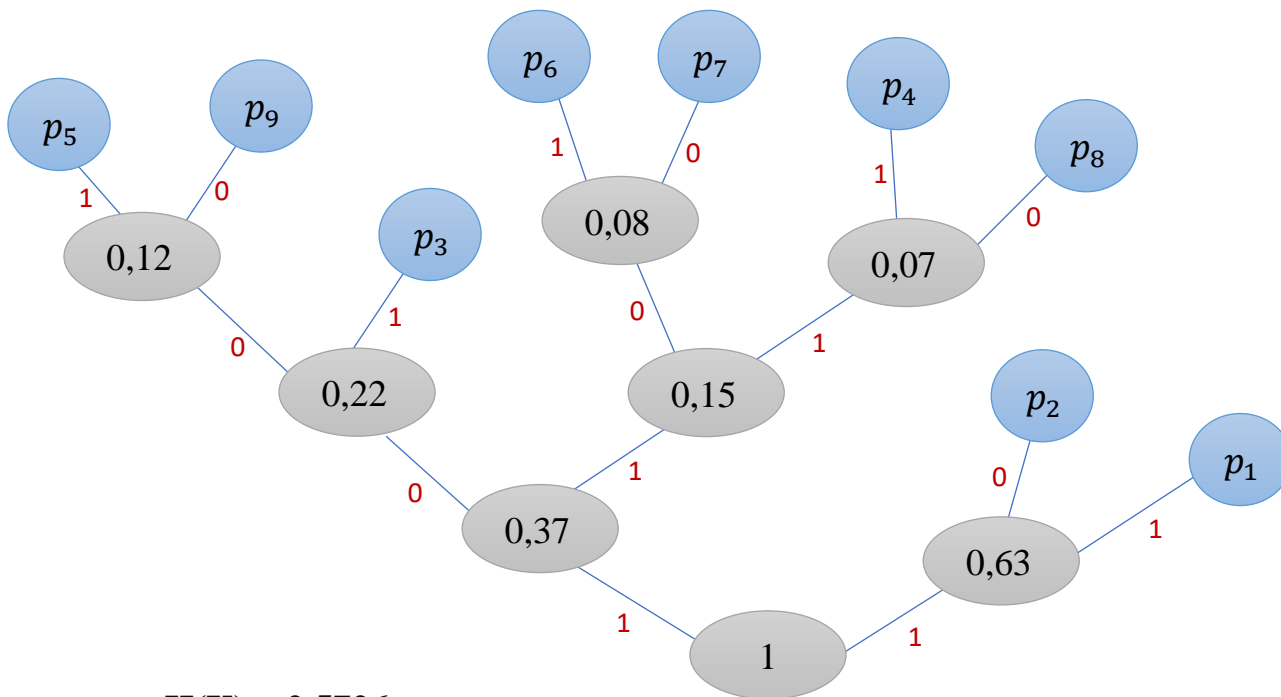
$$H(X) = 2,5726$$

$$\xi = \frac{2,64 - 2,5726}{2,5726} = 0,026$$

Код за алгоритмом Хаффмена:

p_2	0,35	00	2	0,7
p_1	0,28	01	2	0,56
p_3	0,1	101	3	0,3
p_5	0,06	1001	4	0,24
p_9	0,06	1000	4	0,24
p_6	0,04	1101	4	0,16
p_7	0,04	1100	4	0,16
p_8	0,04	1110	4	0,16
p_4	0,03	1111	4	0,12
\tilde{l}				2,64

Дерево Хаффмена:



$$H(X) = 2,5726$$

$$\xi = \frac{2,64 - 2,5726}{2,5726} = 0,026$$

4.2. Побудувала нерівномірні ефективні коди за алгоритмом Шеннона-Фано для кодування поодиноких символів джерела та слів довжиною у два символи. Оцінила та порівняла ефективність отриманих кодів. Побудованими кодами закодувала фрагмент повідомлення довжиною у 30 символів, що був згенерований джерелом

Варіант	p(A)	p(B)	p(C)	p(D)	Фрагмент повідомлення
9	0,46	0,1	0,15	0,29	CCACDACACCCDACACDDBCDCBDABADBD

Код за алгоритмом Шеннона-Фано:

A	0,46	1	1	0,46
D	0,29	01	2	0,58
C	0,15	001	3	0,45
B	0,1	000	3	0,3
\tilde{l}				1,79

$$H(X) = 1,7759$$

$$\xi = \frac{1,79 - 1,7759}{1,7759} = 0,0079$$

Закодований фрагмент:

CCACDACACCCDACACDDBCDCBDABADBD =

00100110010110011001001001011001100101010000010100100001100010100001

L = 68

AA	0,2116	11	2	0,4232
AB	0,1334	101	3	0,4002
BA	0,1334	100	3	0,4002
BB	0,0841	0111	4	0,3364
AC	0,069	0110	4	0,2760
CA	0,069	0101	4	0,2760
AD	0,046	0100	4	0,1840
DA	0,046	0011	4	0,1840
BC	0,0435	00101	5	0,2175
CB	0,0435	00100	5	0,2175
BD	0,029	00011	5	0,1450
DB	0,029	00010	5	0,1450
CC	0,0225	000011	6	0,1350
CD	0,015	000010	6	0,0900
DC	0,015	000001	6	0,0900
DD	0,01	000000	6	0,0600
\tilde{l}				3,58

$$\tilde{l} = 3,58 / 2 = 1,79$$

$$H(X) = 1,7759$$

$$\xi = \frac{1,79 - 1,7759}{1,7759} = 0,0078$$

Закодований фрагмент:

CCACDACACCCDACACDDBCDCBDABADBD =

000011011000110101000011000011011001100000000010100000100011101010000011

$$L = 72$$

4.3. 1. Побудувала нерівномірні ефективні коди за алгоритмом Шеннона-Фано для кодування поодиноких символів джерела та слів довжиною у два символи.

2. Побудувала марковський алгоритм для кодування символів джерела.

3. Оцінила та порівняла ефективність отриманих кодів та марковського алгоритму.

4. Побудованими кодами закодувала фрагмент повідомлення довжиною у 20 символів, що був згенерований джерелом

Варіант	$\begin{pmatrix} P(A A) & P(B A) & P(C A) \\ P(A B) & P(B B) & P(C B) \\ P(A C) & P(B C) & P(C C) \end{pmatrix}$	Фрагмент повідомлення
9	0,3 0,24 0,46 0,44 0,05 0,51 0,47 0,16 0,37	CCBCCABBCABCBVCCCCC

Знаходження безумовних імовірностей виникнення символів:

$$\begin{cases} P(A) = P(A)P(A|A) + P(B)P(A|B) + P(C)P(A|C) \\ P(B) = P(A)P(B|A) + P(B)P(B|B) + P(C)P(B|C) \\ P(A) + P(B) + P(C) = 1 \end{cases}$$

$$\begin{cases} -0,7P(A) + 0,44P(B) + 0,47P(C) = 0 \\ 0,24P(A) - 0,95P(B) + 0,16P(C) = 0 \\ P(A) + P(B) + P(C) = 1 \end{cases}$$

$$\begin{cases} P(A) = 0,397 \\ P(B) = 0,173 \\ P(C) = 0,43 \end{cases}$$

C	0,43	1	1	0,43
A	0,397	01	2	0,794
B	0,173	00	2	0,346
\tilde{l}				1,57

H(X) = 1,4905

$$\xi = \frac{1,57 - 1,4905}{1,4905} = \mathbf{0,053}$$

Закодований фрагмент:

ССВССАВВСАВСВВССССССС = 1100110100001010010000111111

L = 28

AA	$P(A)P(A A) = 0,397 * 0,3 = 0,1191$
AB	$P(A)P(B A) = 0,397 * 0,24 = 0,09528$
AC	$P(A)P(C A) = 0,397 * 0,46 = 0,18262$
BB	$P(B)P(B B) = 0,173 * 0,05 = 0,00865$
BA	$P(B)P(A B) = 0,173 * 0,44 = 0,07612$
BC	$P(B)P(C B) = 0,173 * 0,51 = 0,08823$
CC	$P(C)P(C C) = 0,43 * 0,37 = 0,1591$
CA	$P(C)P(A C) = 0,43 * 0,47 = 0,2021$
CB	$P(C)P(B C) = 0,43 * 0,16 = 0,0688$

CA	0,2021	11	2	0,40420
AC	0,18262	101	3	0,54786
CC	0,1591	100	3	0,47730
AA	0,1191	011	3	0,35730
AB	0,09528	010	3	0,28584
BC	0,08823	001	3	0,26469
BA	0,07612	0001	4	0,30448
CB	0,0688	00001	5	0,34400
BB	0,00865	00000	5	0,04325
\tilde{l}				3,028

$$\tilde{l} = 3,028 / 2 = 1,514$$

$$H(X) = 2,9417$$

$$\xi = \frac{3,028 - 2,9417}{2,9417} = 0,029$$

Закодований фрагмент:

$$CCBCCABBCABCBVCCCCC = 10000111000001100100000100100100$$

$$L = 32$$

Після символу A:

C	0,46	1	1	0,46
A	0,3	01	2	0,6
B	0,24	00	2	0,48
$\tilde{l}_{(A)}$				1,54

Після символу B:

C	0,51	1	1	0,51
A	0,44	01	2	0,88
B	0,05	00	2	0,1
$\tilde{l}_{(B)}$				1,49

Після символу C:

A	0,47	1	1	0,47
C	0,37	01	2	0,74
B	0,16	00	2	0,32
$\tilde{l}_{(C)}$				1,53

$$\tilde{l} = P(A)\tilde{l}_{(A)} + P(B)\tilde{l}_{(B)} + P(C)\tilde{l}_{(C)} = 0,397*1,54 + 0,173*1,49 + 0,43*1,53 = 1,527$$

Закодований фрагмент:

ССВССАВВСАВСВВССССССС = 1010010110100110010000010101010101

L = 34

$$\begin{aligned} H(X|A) &= -(P(A|A) \log_2 P(A|A) + P(B|A) \log_2 P(B|A) + P(C|A) \log_2 P(C|A)) \\ &= -(0,3 \log_2 0,3 + 0,24 \log_2 0,24 + 0,46 \log_2 0,46) = 1,5306 \end{aligned}$$

$$\begin{aligned} H(X|B) &= -(P(A|B) \log_2 P(A|B) + P(B|B) \log_2 P(B|B) + P(C|B) \log_2 P(C|B)) \\ &= -(0,44 \log_2 0,44 + 0,05 \log_2 0,05 + 0,51 \log_2 0,51) = 1,2327 \end{aligned}$$

$$\begin{aligned} H(X|C) &= -(P(A|C) \log_2 P(A|C) + P(B|C) \log_2 P(B|C) + P(C|C) \log_2 P(C|C)) \\ &= -(0,47 \log_2 0,47 + 0,16 \log_2 0,16 + 0,46 \log_2 0,37) = 1,4503 \end{aligned}$$

$$\begin{aligned} H(X) &= P(A)H(X|A) + P(B)H(X|B) + P(C)H(X|C) \\ &= 0,397 * 1,5306 + 0,173 * 1,2327 + 0,43 * 1,4503 = 1,4445 \end{aligned}$$

$$\xi = \frac{1,527 - 1,4445}{1,4445} = \mathbf{0,057}$$