# КРИПТОГРАФІЯ КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Експериментальна оцінка ентропії на символ джерела відкритого тексту

Виконали: ФБ-31 Голомовза Дар`я ФБ-31 Караман Любов

#### Мета роботи:

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

#### Порядок виконання роботи

- 0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
- 1. Написати програми для підрахунку а) частот букв і b) частот біграм в тексті, а також підрахунку с) H1 та H2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H1 та H2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення H1 та H2 на тому ж тексті, в якому вилучено всі пробіли.
- А) Частота букв

### Частота букв рахується за формулою:

Частота(символ) = 
$$\frac{\text{Кількість появ символу}}{\text{Загальна кількість символів}} = \frac{f_i}{N}$$

Створюємо код який автоматично здійснить підрахунки та збереже результат до окремого файлу:

[Running] python -u "d:\3 kypc\Crypto\crypto25-26\tasks\cp1\lab1.py"
Result saved to crypto\_analysis.xlsx

Частота літер з пробілами:

Літера	Частота
	0,160251
а	0,072611
б	0,013002
В	0,039583
Γ	0,016126
Д	0,023597
е	0,068249
ж	0,007564
3	0,014923
И	0,057457
й	0,009841
К	0,030708
Л	0,044157
М	0,025196
Н	0,05383
0	0,092891
П	0,02387
р	0,039848
С	0,042555
Т	0,050624
У	0,025231
ф	0,001812

х	0,006909
ц	0,002785
Ч	0,013351
Ш	0,00733
щ	0,00297
ъ	0,000269
Ы	0,01452
Ь	0,014918
Э	0,00258
ю	0,004433
Я	0,016011

## Без пробілу:

•	•
Літера	Частота
а	0,086468
б	0,015483
В	0,047137
Г	0,019203
Д	0,028101
е	0,081273
ж	0,009007
3	0,017771
И	0,068421
й	0,011719
К	0,036568
Л	0,052584
М	0,030004
Н	0,064102
0	0,110617
П	0,028425
р	0,047453
С	0,050675
Т	0,060284
У	0,030046
ф	0,002158
Х	0,008227
ц	0,003316
Ч	0,015899
Ш	0,008728
щ	0,003537
ъ	0,000321
Ы	0,017291
Ь	0,017764
Э	0,003072
Ю	0,005278
Я	0,019066

## Біграми які перетинаються з пробілами

		а	6	В	г	д	e	ж	3	и	й	к	л	M	н	0	п	р	С	т	у	φ	×	ц	ч	ш	щ	ъ	ы	ь	9	ю	я
	0,0046	0,0028	0,0063	0,016	0,0044	0,0053	0,0028	0,0018	0,0046	0,0109	0	0,0094	0,0027	0,0058	0,0148	0,0095	0,0166	0,0039	0,0146	0,0072	0,004	0,0007	0,0011	0,0003	0,0051	0,001	0,0001	0	0	0	0,0024	0,0001	0,0016
a	0,0175	0	0,0006	0,0025	0,0009	0,0022	0,0011	0,0011	0,0039	0,0003	0,0006	0,0043	0,0086	0,0031	0,0051	0,0001	0,0009	0,0044	0,004	0,0049	0,0001	0,0002	0,0011	0,0001	0,0009	0,0008	0,0003	0	0	0	0	0,0007	0,0021
6	0,0002	0,001	0	0	0	0	0,0023	0	0	0,0007	0	0,0002	0,0006	0	0,0003	0,0019	0	0,0011	0,0001	0	0,0012	0	0	0	0	0	0,0001	0,0001	0,0027	0	0	0	0,0003
В	0,0068	0,0063	0	0,0001	0	0,0003	0,0056	0	0,0004	0,0035	0	0,0002	0,0006	0,0001	0,0009	0,0063	0,0002	0,0006	0,0021	0,0003	0,0006	0	0	0	0,0001	0,001	0	0	0,0026	0,0005	0	0	0,0003
r	0,0005	0,0021	0	0	0	0,0009	0,0005	0	0	0,0006	0	0,0001	0,0016	0	0,0004	0,0076	0	0,0013	0	0	0,0005	0	0	0	0	0	0	0	0	0	0	0	0
А	0,0014	0,0042	0,0001	0,0009	0	0	0,0038	0,0001	0	0,002	0	0,0003	0,0004	0,0001	0,0019	0,0033	0,0001	0,001	0,0004	0,0002	0,0015	0	0	0,0002	0,0001	0,0002	0	0,0001	0,0005	0,0005	0	0	0,0004
e	0,0158	0,0001	0,0009	0,0021	0,0029	0,0028	0,0013	0,0008	0,0014	0,0002	0,0019	0,0015	0,0059	0,0037	0,0063	0,0002	0,0012	0,0065	0,0041	0,0052	0,0002	0,0001	0,0004	0,0003	0,0011	0,0006	0,0006	0	0	0	0	0,0002	0,0002
ж	0,0003	0,0013	0	0	0	0,0007	0,003	0	0	0,0011	0	0,0001	0	0	0,0006	0,0001	0	0	0	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	0
3	0,0016	0,0055	0,0002	0,0011	0,0003	0,0009	0,0006	0,0001	0,0001	0,0005	0	0,0001	0,0003	0,0002	0,0014	0,0006	0	0,0003	0	0	0,0004	0	0	0	0	0	0	0	0,0004	0,0001	0	0	0,0002
И	0,016	0,0002	0,0006	0,0028	0,0005	0,0015	0,0017	0,0002	0,0022	0,0005	0,0016	0,0027	0,0061	0,0025	0,0033	0,0004	0,0003	0,0011	0,0029	0,0041	0,0001	0,0001	0,0014	0,001	0,0017	0,0005	0,0002	0	0	0	0	0,0002	0,001
й	0,0078	0	0	0	0	0,0001	0	0	0	0	0	0,0001	0	0	0,0004	0	0	0	0,0004	0,0004	0	0	0	0,0001	0,0002	0,0002	0	0	0	0	0	0	0
K	0,0042	0,0071		0,0005	0	0	0,0009	0	0	0,0029	0	0	0,0008	0	0,0007	0,0084	0	0,0021	0,0001	0,0007	0,0022	0	0	0	0	0	0	0	0	0	0	0	0
л	0,0078	0,0075	0,0001	0	0,0001	0,0001		0,0002	0	0,0062	0	0,0004	0,0004	0,0001	0,0005	0,0062	0,0001	0	0,002	0,0002	0,0013	0	0	0	0,0001	0	0	0	0,0009	0,0035	0	0,0007	0,0013
M	0,0072	0,0036	0	0,0001	0,0001	0	0,0031	0	0	0,0028	0	0,0001	0,0002	0,0001	0,0013	0,0029	0,0002	0,0001	0,0003	0,0001	0,0017	0	0	0	0	0	0	0	0,0007	0	0	0	0,0005
н	0,0033	0,0091	0	0	0,0001	0,0009	0,0088	0	0,0001	0,007	0	0,0003	0	0	0,0031	0,0097	0	0	0,0005	0,0008	0,0037	0,0001	0	0,0003	0,0002	0	0,0002	0	0,0032	0,0009	0	0,0002	0,0014
0	0,0207	0	0,003	0,007	0,0043	0,0045	0,0015	0,0016	0,0015	0,0007	0,0038	0,0026	0,0063	0,0057	0,0054	0,0002	0,0014	0,0062	0,0056	0,0065	0,0001	0,0003	0,0005	0,0001	0,0017	0,0008	0,0002	0	0	0	0,0001	0,0004	0,0005
п	0	0,0017	0	0	0	0	0,0021	0	0	0,0013	0	0,0001	0,0009	0	0,0003	0,0092	0,0001	0,0065	0	0,0002	0,0008	0	0	0	0	0	0	0	0,0003	0,0001	0	0	0,0004
р	0,0014	0,0075		0,0005	0,001	0,0003	0,0056	0,0002	0	0,0056	0	0,0004	0,0003	0,0002	0,001	0,0079	0,0001	0,0001	0,0002	0,0009	0,0025	0	0,0002	0,0001	0,0002	0,0004	0	0	0,0013	0,0006	0	0,0002	0,001
C	0,0029	0,0017	0,0001	0,0015	0	0,0003	0,0029	0	0	0,0018	0	0,0039	0,0025	0,0006	0,0009	0,0026	0,0016	0,0002	0,0009	0,0092	0,0007	0	0,0001	0,0001	0,0003	0,0001	0	0	0,0003	0,0034	0	0,0001	0,0037
T	0,0056	0,0056	0	0,002	0	0,0002	0,0057	0	0	0,0034	0	0,0007	0,0002	0	0,0013	0,0132	0,0001	0,0029	0,0008	0,0001	0,002	0,0001	0	0	0,0004	0	0	0	0,0014	0,0045	0	0,0001	0,0004
У	0,0064	0,0001	0,0006	0,0007	0,0009	0,0018		0,0012	0,0003	0	0,0001	0,001	0,0022	0,0009	0,0005	0	0,0006	0,001	0,0012	0,0016	0	0,0002	0,0008	0	0,0007	0,0007	0,0002	0	0	0	0	0,001	0,0001
ф	0	0,0003	0	0	0	0	0,0005	0	0	0,0003	0	0	0,0001	0	0	0,0002	0	0,0002	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	0
X	0,0029	0,0006		0,0002	0,0001	0	0,0001	0	0	0,0003	0	0	0,0002	0,0001	0,0003	0,0018	0	0,0002	0	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	0
ц	0,0003	0,0006	0	0,0001	0	0	0,0008	0	0	0,0003	0	0	0	0	0	0,0003	0	0	0	0	0,0002	0	0	0	0	0	0	0	0,0002	0	0	0	0
ч	0,0005	0,0022	0	0	0	0	0,0038	0	0	0,0015	0	0,0005	0	0	0,0008	0,0001	0	0,0001	0	0,003	0,0005	0	0	0	0	0,0001	0	0	0	0,0003	0	0	0
ш	0,0001	0,001	0	0,0001	0	0	0,0021	0	0	0,0018	0	0,0004	0,0004	0	0,0004	0,0002	0	0	0	0,0001	0,0004	0	0	0	0	0	0	0	0	0,0003	0	0	0
щ	0	0,0004	0	0	0	0	0,0012	0	0	0,001	0	0	0	0	0,0001	0	0	0	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	
ъ	0	0	0	0	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0002
ы	0,0037	0	0,0003	0,0008	0,0001	0,0001	0,0009	0,0001	0,0001	0	0,0018	0,0002	0,0019	0,0012	0,0002	0	0,0001	0,0003	0,0007	0,0006	0	0	0,0008	0	0,0002		0	0	0	0	0	0	0
ь	0,0091	0	0	0,0001	0,0001	0,0001	0,0008	0	0,0001	0,0001	0	0,0009	0	0,0002	0,0013	0	0	0	0,0007	0,0003	0	0	0	0,0001	0,0001	0,0003	0	0	0	0	0	0,0003	0,0003
,	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0024	0	0	0	0	0	0	0	0	0	0	0	0	0
ю	0,0025	0	0,0002	0	0	0,0002	0	0	0	0	0	0	0	0,0001	0	0	0	0,0001	0,0002	0,0003	0	0	0,0001	0	0,0001	0	0,0005	0	0	0	0	0,0001	0
Я	0,0092	0	0	0,0005	0,0001	0,0006	0,0001	0,0001	0,0003	0	0	0,0002	0,0008	0,0003	0,0006	0	0,0001	0,0001	0,0008	0,0013	0	0	0,0001	0	0,0002	0	0,0004	0	0	0	0	0,0001	0,0001

## Біграми які не перетинаються з пробілами

		a	6	8	r	Д	c	ж	3	И	й	K	л	м	н	0	п	р	c	т	у	ф	×	ц	ч	w	щ	ъ	ы	b	9	ю	Я
	0,0046	0,0028	0,0062	0,016		0,0052	0,0028	0,0019	0,0045	0,0109	0	0,0093	0,0027	0,0059	0,0148	0,0095	0,0167	0,0038	0,0146	0,0073	0,0041	0,0007	0,0011	0,0003	0,0052	0,001	0,0002	0	0	0	0,0024	0	0,0015
a	0,0172	0	0,0006	0,0025	0,0009	0,0023	0,0011	0,0011	0,0037	0,0003	0,0006	0,0042	0,0086	0,0031	0,0051	0,0001	0,0009	0,0043	0,0042	0,0049	0,0001	0,0002	0,0011	0,0001	0,001	0,0008	0,0004	0	0	0	0	0,0008	
6	0,0002	0,001	0	0	0	0	0,0023	0	0	0,0007	0	0,0002	0,0006	0	0,0003	0,002	0	0,001	0,0001	0	0,0012	0	0	0	0	0	0,0001	0,0001	0,0028	0	0	0	0,0003
8	0,0068	0,0063	0	0,0001	0	0,0002	0,0056	0	0,0004	0,0035	0	0,0002	0,0006	0,0001	0,0009	0,0063	0,0002		0,0021	0,0003	0,0007	0	0,0001	0	0,0001	0,001	0	- 0	0,0025	0,0005	0	0	0,0003
г	0,0005	0,0022	0	0	0	0,0009	0,0004	0	0	0,0006	0	0,0001	0,0014	0	0,0004	0,0076	0	0,0013	0	0	0,0005	0	0	0	0	0	0	0	0	0	0	0	0
Д	0,0014	0,0042	0,0001	0,0009	0	0	0,0038	0,0001	0	0,0019	0	0,0003	0,0004	0,0001	0,0019	0,0033	0,0001	0,0009	0,0004	0,0002	0,0015	0	0	0,0002	0,0001	0,0002	0	0,0001	0,0005	0,0005	0	0	0,0004
e	0,0159	0,0001	0,0009	0,0022	0,003	0,0029	0,0012	0,0007	0,0013	0,0002	0,0018	0,0015	0,0059	0,0038	0,0063	0,0002	0,0012	0,0065	0,004	0,0053	0,0002	0,0001	0,0004	0,0003	0,001	0,0007	0,0006	0	0	0	0	0,0002	0,0002
ж	0,0003	0,0013	0	0	0	0,0007	0,0029	0	0	0,0011	0	0,0001	0	0	0,0006	0,0001	0	0	0	0	0,0002	0	0	0	0	0	0	0	0	0	0	. 0	0
3	0,0016	0,0056		0,0011		0,0009	0,0007	0,0001		0,0005	0	0,0001	0,0004	0,0002	0,0014	0,0007	0,0001	0,0003	0	0	0,0005	0	0	0	0	0	0	0	0,0004	0,0001	0	0	0,0002
и	0,0158	0,0002	0,0005	0,0029	0,0005	0,0015	0,0017	0,0002	0,0021	0,0005	0,0016		0,0062	0,0025	0,0032	0,0004	0,0003	0,0011	0,0029		0,0001	0,0001	0,0014	0,001	0,0017	0,0005	0,0002	0	0	0	0	0,0002	0,0009
й	0,0077	0	0	0	0	0,0001	0	0	0	0	0	0,0002	0	0,0001	0,0004	0	0	0	0,0004	0,0004	0	0	0	0,0001	0,0002	0,0002	0	0	0	0	0	. 0	0
K	0,0041	0,0072	0	0,0006		0	0,0008	0,0001	0	0,0029	0	0	0,0008	0	0,0007	0,0085	0	0,0021	0,0001	0,0007	0,0022	0	0	0	0	0	0	0	0	0	0	0	0
л	0,0079	0,0075	0,0001	0	0,0001	0,0001	0,0045	0,0002	0	0,0061	0	0,0004	0,0005	0,0001	0,0005	0,0063	0,0001	0	0,0019	0,0002	0,0012	0	0	0	0,0001	0	0	0	0,0008	0,0034	0	0,0007	
M	0,0071	0,0036	0	0,0001		0	0,0031	0	0	0,0029	0	0,0001	0,0002	0,0001	0,0013	0,0028	0,0002	0,0001	0,0002	0,0001	0,0017	0	0	0	0	0	0	0	0,0007	0	0	0	0,0005
н	0,0032	0,0091	0	0	0,0001	0,0009	0,0088	0	0	0,0071	0	0,0003	0	0	0,0031	0,0097	0	0,0001	0,0005	0,0008	0,0036	0,0001	0	0,0003	0,0002	0	0,0002	0	0,0032	0,0009	0	0,0001	
0	0,0206	0	0,0029	0,0069	0,0044	0,0045		0,0017	0,0015	0,0007	0,0038	0,0025	0,0062	0,0058	0,0053	0,0002	0,0013		0,0055	0,0064	0,0001	0,0003	0,0005	0,0001	0,0017	0,0007	0,0003	0	0	0	0,0001	0,0004	
n	0	0,0017	0	0	0	0	0,0021	0	0	0,0012	0	0,0001	0,0009	0	0,0003	0,0091	0,0001	0,0063	0	0,0002	0,0008	0	0	0	0	0	0	0	0,0003	0,0001	0	0	0,0004
p	0,0015	0,0075	0,0001	0,0005	0,001	0,0002	0,0057	0,0002	0	0,0057	0	0,0004	0,0003	0,0002	0,001	0,0079	0,0001	0,0001	0,0002	0,0009	0,0024	0	0,0002	0,0001	0,0002	0,0004	0	0	0,0013	0,0006	0	0,0002	0,001
c	0,0029	0,0016	0,0001	0,0014	0	0,0003	0,0029	0	0	0,0018	0	0,0039	0,0026	0,0006	0,0009	0,0026	0,0017	0,0002	0,0009	0,0092	0,0007	0	0,0002	0,0001	0,0003	0,0001	0	0	0,0003	0,0034	0	0,0001	0,0038
т	0,0057	0,0055	0	0,002	0	0,0002	0,0056	0	0	0,0034	0	0,0007	0,0002	0	0,0013	0,0133	0,0001	0,0029	0,0008	0,0001	0,002	0,0001	0	0	0,0004	0	0	0	0,0014	0,0044	0	0,0001	0,0005
У	0,0063	0,0001	0,0007	0,0007	0,0009	0,0019	0,0002	0,0012	0,0003	0	0,0001	0,001	0,0023	0,001	0,0005	0	0,0006	0,0009	0,0012	0,0017	0	0,0002	0,0008	0	0,0007	0,0006	0,0002	0	0	0	0	0,001	0,0001
ф	0	0,0003	0	0	0	0	0,0005	0	0	0,0003	0	0	0,0001	0	0	0,0003	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
×	0,0029	0,0006	0	0,0002	0,0001	0	0,0001	0	0	0,0003	0	0	0,0002	0	0,0003	0,0017	0	0,0002	0,0001	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	0
ц	0,0003	0,0007	0	0,0001	0	0	0,0008	0	0	0,0003	0	0	0	0	0	0,0003	0	0	0	0	0,0001	0	0	0	0	0	0	0	0,0002	0	0	0	0
ч	0,0006	0,0021	0	0	0	0	0,0038	0	0	0,0016	0	0,0005	0	0	0,0008	0,0001	0	0	. 0	0,003	0,0005	0	0	0	0	0	0	0	0	0,0003	0	0	0
ш	0,0001	0,001	0	0,0001	0	0	0,0021	0	0	0,0018	0	0,0004	0,0005	0	0,0004	0,0002	0,0001	0	0	0,0001	0,0004	0	0	0	0	0	0	0	0	0,0003	0	0	0
ш	0	0,0004	0	0	0	0	0,0012	0	0	0,001	0	0	0	0	0,0001	0	0	0	0	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	0
ъ	0	0	0	0	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0002
ы	0,0038	0	0,0003	0,0008	0,0001	0,0001	0,0009	0,0001	0,0001	0	0,0018	0,0002	0,0018	0,0012	0,0002	0	0,0002	0,0002	0,0007	0,0005	0	0	0,0008	0	0,0002	0,0005	0	0	0	0	0	0	0
ь	0,0092	0	0	0,0001	0,0001	0,0001	0,0008	0	0,0001	0,0001	0	0,0009	0	0,0002	0,0013	0	0	0	0,0007	0,0002	0	0	0	0,0001	0,0001	0,0003	0	0	0	0	0	0,0003	0,0002
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0024	0	0	0	0	0	0	0	0	0	0	0		0
ю	0,0026	0	0,0001	0	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	0	0,0002	0,0003	0	0	0,0001	0	0,0001	0	0,0004	0	0	0	0	0,0001	0
Я	0,0092	0	0	0,0005	0,0001	0,0006	0,0001	0,0001	0,0003	0,0001	0	0,0002	0,0009	0,0003	0,0006	0	0,0001	0,0001	0,0008	0,0013	0	0	0,0002	0	0,0002	0	0,0004	0	0	0	0	0,0001	0,0001

## Біграми які перетинаються без пробілів

		-	U	-		u	- 11	-	,		-	IVI	IV	0	-	ų	n .	3		U	- 1	vv	^		-	AN	AD	AL	AU	ML	Al	MO
_	a	6	В	r	. д	e	ж	3	И	Й	К	л	M	н	0	n	р	С	T	У	Φ	X	ц	ч	ш	щ	ъ	ы	ь	9	ю	я
a	0,0005	0,0014	0,0051	0,0017	0,0034	0,0017	0,0015	0,0052	0,0018	0,0007	0,0064	0,0107	0,0046		0,0012	0,0034	0,0057	0,0068	0,0069		0,0003	0,0015	0,0001	0,0016	0,0011	0,0004	0	0	0	0,0004	0,0009	
6	0,0011	0	0	0	0	0,0028	0	0	0,0008	0	0,0002		0,0001		0,0023	0	0,0013	0,0001	0	0,0015	0	0	0	0	0	0,0002	0,0002	0,0033	0	0,0001	- 0	0,0004
8	0,0076	0,0003	0,0006	0,0004	0,0006	0,0068	0,0001	0,0006	0,0045	0	0,0011		0,0005		0,008	0,0011	0,0011	0,0033	0,001		0	0,0001	0,0001	0,0003	0,0012	0	0	0,0031	0,0006	0,0003	0	0,0004
Г	0,0026	0	0,0001	0	0,0011	0,0006	0	0	0,0008	0	0,0001	0,0019	0	0,0005	0,0091	0,0001	0,0015	0,0001		0,0006	0	0	0	0,0001	0	0	0	0	0	0	0	₩-
А	0,005	0,0001	0,0012	0,0001	0,0001	0,0046	0,0001	0,0001	0,0025	0	0,0005		0,0002	0,0025	0,0039	0,0003	0,0012	0,0006	0,0003	0,0018	0	0	0,0003	0,0001	0,0002	0	0,0001	0,0006	0,0006	0		0,0009
e	0,0004	0,002	0,0044	0,004	0,004	0,0018	0,0011	0,0023	0,0015	0,0022	0,0026		0,0052	0,0091	0,0014	0,0036	0,0082	0,0066	0,0071	0,0008	0,0002	0,0007	0,0004	0,0018	0,0009	0,0007	0	0	0	0,0003	0,0002	0,0004
ж	0,0016	0	0	0	0,0009	0,0035	0	0	0,0014	0	0,0002	0,0001	0	0,0008	0,0001	0	0	0	0	0,0002	0	0	0	0	0	0	0	0	0	0		
3	0,0065	0,0003	0,0015	0,0004	0,0011	0,0008	0,0001	0,0002	0,0007	0	0,0003		0,0004		0,0009	0,0002	0,0004	0,0002	0,0001	0,0005	0	0	0	0,0001	0	0	0	0,0005	0,0002	0	0	0,0003
И	0,0005	0,0014	0,0056	0,0012	0,0025	0,0023	0,0004	0,0033	0,0017	0,0019	0,0042		0,0036		0,0017	0,0026	0,0018	0,0054	0,0058	0,0007	0,0002	0,0018	0,0012	0,0026	0,0007	0,0003	0	0	0	0,0002	0,0003	
й	0,0002	0,0004	0,0008	0,0004	0,0005	0,0001	0,0002	0,0002	0,0007	0	0,0008		0,0005		0,0005	0,0011	0,0003	0,0014	0,0008	0,0002	0,0001	0,0001	0,0001	0,0005	0,0003	0	0	0	0	0,0001		0,0001
K	0,0085	0,0004	0,0012	0,0001	0,0001	0,0011	0,0002	0,0001	0,0038	0	0,0003		0,0002		0,0103	0,0004	0,0026	0,0007	0,0011	0,0027	0	0	0	0,0002	0	0	0	0	0	0,0001	0	0,0001
л	0,0093	0,0005	0,0011	0,0004	0,0003	0,0056	0,0003	0,0002	0,0082	0	0,0011		0,0004		0,0079	0,001	0,0003	0,0031	0,0006		0,0001	0	0	0,0006	0,0001	0	0	0,001	0,0042		0,0008	
м	0,0044	0,0004	0,0009	0,0005	0,0003	0,0038	0,0001	0,0003	0,004	0	0,0006		0,0004		0,004	0,0011	0,0003	0,0012	0,0004	0,0023	0	0,0001	0	0,0005	0,0001	0	0	0,0008	0,0001	0,0001	0	0,0007
н	0,011	0,0002	0,0004	0,0002	0,0011	0,0105	0,0001	0,0002	0,0087	0	0,0005		0,0001	0,0041	0,0118	0,0005	0,0001	0,001	0,0011	0,0046	0,0001	0	0,0004	0,0003	0	0,0002	0	0,0038	0,0011	0	0,0002	
0	0,0003	0,0047	0,0108	0,0057	0,0062	0,0023	0,0023	0,0025	0,0022	0,0045	0,0042		0,0076		0,0021	0,0041	0,0079	0,009	0,0089	0,0007	0,0004	0,0007	0,0002	0,003	0,0011	0,0003	0	0	0	0,0006	0,0005	
п	0,002	0	0	0	0	0,0025	0	0	0,0015	0	0,0001	0,0011	0	0,0003	0,0109	0,0001	0,0077	0	0,0002	0,0009	0	0	0	0	0	0	0	0,0004	0,0001	0		0,0005
р	0,009		0,0007	0,0012	0,0004	0,0067	0,0003	0,0001	0,0069	0	0,0006	0,0003			0,0095	0,0003	0,0001	0,0003		0,003	0	0,0003	0,0001	0,0003		0	0	0,0016	0,0007	0	0,0003	
c	0,002	0,0002	0,0021	0,0002	0,0004	0,0036	0,0001	0,0001	0,0023	0	0,0049		0,0008	0,0014	0,0033	0,0023	0,0003	0,0014		0,001	0		0,0001	0,0005	0,0002	0	0	0,0004	0,0041		0,0001	
т	0,0068	0,0003	0,003	0,0002	0,0004	0,0069	0,0002	0,0002	0,0046	0	0,0013		0,0003	0,0021	0,0162	0,0008	0,0035	0,0015	0,0004	0,0025	0,0001	0,0001	0	0,0007	0	0	0	0,0017	0,0054		0,0001	
у	0,0003	0,001	0,0017	0,0012	0,0024	0,0003	0,0015	0,0005	0,0008	0,0002	0,0016	0,0027	0,0014	0,0015	0,0004	0,0015	0,0013	0,002	0,0023	0,0002	0,0003	0,001	0	0,0012	0,0009	0,0003	0	0	0	0,0001	0,0012	0,0002
ф	0,0003	0	0	0	0	0,0006	0	0	0,0004	0	0	0,0001	0	0	0,0003	0	0,0002	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	0
X	0,0008	0,0001	0,0005	0,0002	0,0002	0,0001	0	0,0001	0,0007	0	0,0003	0,0003	0,0002	0,0007	0,0023	0,0004	0,0003	0,0004	0,0002	0,0003	0	0	0	0,0001	0,0001	0	0	0	0	0		-0
ц	0,0008	0	0,0001	0	0	0,001	0	. 0	0,0004	0	0	0	0	0	0,0003	0	0,0001	0	. 0	0,0002	0	0	0	0	0	0	0	0,0002	0	0	0	100
ч	0,0027	0	0,0001	0	0	0,0045	0	0	0,0019	0	0,0006	0	0	0,001	0,0001	0,0001	0,0001	0,0001	0,0036	0,0006	0	0	0	0	0,0001	0	0	0	0,0003	0	0	- 0
ш	0,0012	0	0,0001	0	0	0,0024	0	0	0,0022	0	0,0005	0,0005	0	0,0005	0,0003	0,0001	0	0	0,0001	0,0005	0	0	0	0	0	0	0	0	0,0004	0	0	- 0
щ	0,0005	0	0	0	0	0,0015	0	0	0,0012	0	0	0	0	0,0001	0	0	0	0	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	0
ъ	0	0	0	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0,0002
ы	0,0001	0,0005	0,0015	0,0002	0,0003	0,0012	0,0001	0,0002	0,0004	0,0021	0,0005		0,0016	0,0007	0,0003	0,0006	0,0004	0,0012	0,0009	0,0002	0	0,001	0	0,0003	0,0006	0	0	0	0	0,0001	0	0,0001
ь	0,0002	0,0004	0,0013	0,0003	0,0005	0,0012	0	0,0004	0,0009	0	0,0017	0,0002	0,0007	0,0027	0,0008	0,0011	0,0002	0,0018	0,0009	0,0002	0,0001	0,0001	0,0002	0,0006	0,0004	0	0	0	0	0,0001	0,0004	0,0004
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0029	0	0	0	0	0	0	0	0	0	0	0	0	-
ю	0,0001	0,0003	0,0003	0,0001	0,0004		0,0001	0,0001	0,0002	0	0,0003		0,0002		0,0002	0,0003	0,0002	0,0005	0,0005		0	0,0001	0	0,0003		0,0006	0	0	0	0	0,0001	(
я	0,0002	0,0004	0,0019	0,0004	0,001	0,0003	0,0003	0,0006	0,0009	0,0001	0,001	0,0011	0,0007	0,0019	0,0006	0,0011	0,0003	0,0019	0,002	0,0003	0,0001	0,0003	0,0001	0,0005	0,0001	0,0004	0	0	0	0,0001	0,0002	0,0002

## Біграми які не перетинаються з без пробілів

	a	6	В		п	c	ж	3	и	й	К	л	M	н	0	п	D	c	т .	v	ф	×	ш	ч	ш	ш	ъ	ы	ь	,	ю	8
a	0.0005	0.0014	0.0052	0.0017	0.0034	0.0019	0.0014	0.0049	0.0018	0.0007	0.0065	0.0105	0.0046	0.0084	0.0012	0.0034	0.0056	0.0069	0.0067	0.0007	0.0003	0.0014	0.0001	0.0016	0.0012	0.0004	0	. 0	0	0.0003	0.0008	
6	0.0011	0	0	0	0	0.0029	0	0	0.0008	0	0.0002	0.0007	0	0.0004	0.0023	0	0.0013	0.0001	0	0.0015	0	0	0	0	0	0.0002	0.0002	0.0031	0	0.0001	0	0.0004
В	0,0075	0,0003	0,0007	0,0004	0,0006	0,0068	0,0001	0,0006	0,0044	0	0,0012	0,0009	0,0005	0,0016	0,008	0,001	0,0011	0,0033	0,0011	0,0009	0	0,0001	0	0,0003	0,0012	0	0	0,0031	0,0005	0,0003	0	0,000
r	0,0025	0	0,0001	0	0,0011	0,0006	0	0	0,0009	0	0,0001	0,0018	0	0,0005	0,009	0,0001	0,0016	0,0001	0	0,0006	0	0	0	0,0001	0	0	0	0	0	0	0	, ,
Д	0,0052	0,0001	0,0012	0,0001	0	0,0046	0,0001	0,0001	0,0024	0	0,0005	0,0005	0,0002	0,0025	0,0039	0,0004	0,0012	0,0006	0,0003	0,0018	0	0	0,0003	0,0001	0,0002	0	0,0001	0,0006	0,0006	0	0	0,000
e	0,0005	0,002	0,0043	0,0039		0,0018	0,0012	0,0023	0,0015	0,0022	0,0026	0,0075	0,0053		0,0014	0,0036	0,008	0,0065	0,0071	0,0008	0,0002	0,0007	0,0004	0,0018	0,0009	0,0007	0	0	0	0,0003	0,0002	0,000
ж	0,0015	0	0	0	0,0008	0,0034	0	0	0,0013	0	0,0002	0,0001	0	0,0008	0,0001	0	0	0	0	0,0002	0	0	0	0,0001	0	0	0	0	0	0	0	
3	0,0066	0,0003	0,0015	0,0005	0,0012	0,0008	0,0001	0,0002	0,0006	0	0,0003	0,0004	0,0004	0,0019	0,0009	0,0002	0,0003	0,0002	0,0001	0,0005	0	0	0	0,0001	0	0	0	0,0004	0,0002	0	0	0,000
И	0,0006	0,0014	0,0056	0,0012	0,0025	0,0024	0,0005	0,0033	0,0017	0,0019	0,0042	0,0075	0,0037	0,0057	0,0017	0,0026	0,0018	0,0054	0,006	0,0007	0,0002	0,0018	0,0012	0,0025	0,0008	0,0002	0	0	0	0,0003	0,0003	0,001
й	0,0002	0,0005	0,0008	0,0005		0,0001	0,0002		0,0007	0	0,0008	0,0002	0,0004	0,0011	0,0005	0,0011	0,0003	0,0014	0,0008	0,0002	0,0001	0,0001	0,0001	0,0005	0,0003	0	0	0	0	0,0001		0,000
К	0,0085	0,0004	0,0011	0,0001		0,001	0,0002		0,0038	0	0,0003	0,0011	0,0002		0,0099	0,0004	0,0025	0,0007	0,0011	0,0027	0	0	0	0,0002	0,0001	0	0	0	0	0,0001	0	0,000
л	0,0092	0,0005	0,0012	0,0005	0,0003	0,0057	0,0003	0,0002	0,0081	0	0,0012	0,0006	0,0003	0,0014	0,0079	0,001	0,0003	0,0032	0,0006	0,0018	0,0001	0	0	0,0006	0,0001	0	0	0,0011	0,0042	0,0001	0,0008	0,0016
м	0,0044	0,0003	0,0009	0,0005	0,0003	0,0038	0,0001	0,0002	0,0039	0	0,0006	0,0005	0,0004	0,0022	0,0041	0,0011	0,0003	0,0012	0,0004	0,0022	0,0001	0,0001	0	0,0005	0,0001	0	0	0,0009	0,0001	0,0001	0	0,000
н	0,0109	0,0002	0,0005	0,0002		0,0104	0,0001	0,0002	0,0087	0	0,0005	0	0,0001	0,0039	0,012	0,0005	0,0001	0,001	0,0012	0,0047	0,0001	0	0,0004	0,0003	0	0,0002	0	0,0039	0,0011	0	0,0002	0,001
0	0,0003	0,0048	0,011	0,0058	0,0063	0,0023	0,0023	0,0024	0,0022	0,0047	0,0043	0,0079	0,0076		0,0021	0,0041	0,0078	0,0092	0,0089	0,0007	0,0004	0,0007	0,0002	0,0031	0,0011	0,0003	0	0	0	0,0005	0,0005	0,000
п	0,0021	0	0	0	0	0,0025	0	0	0,0016	0	0,0001	0,0011	0	0,0003	0,0107	0,0001	0,0079	0	0,0002	0,001	0	0	0	0	0	0	0	0,0004	0,0001	0	0	0,0004
р	0,0092	0,0001	0,0008	0,0012		0,0065	0,0003		0,007	0	0,0005	0,0003	0,0003		0,0094	0,0003	0,0001	0,0003	0,0012	0,003	0	0,0003	0,0001	0,0003	0,0005	0	0	0,0016	0,0008	0	0,0003	0,001
c	0,002	0,0002	0,0021	0,0002	0,0004	0,0036	0,0002	0,0001	0,0021	0	0,0051	0,003	0,0008	0,0014	0,0032	0,0023	0,0003	0,0014	0,0112	0,0009	0,0001	0,0002	0,0001	0,0005	0,0002	0	0	0,0004	0,0041	0,0001	0,0001	0,004
T	0,0068	0,0002	0,0029	0,0002	0,0004	0,0069	0,0002	0,0002	0,0045	0	0,0013	0,0004	0,0003	0,0021	0,016	0,0008	0,0035	0,0015	0,0004	0,0025	0,0001	0,0001	0	0,0008	0	0	0	0,0017	0,0054	0,0001	0,0001	0,000
y	0,0003	0,0009	0,0017	0,0012	0,0024	0,0003	0,0015	0,0005	0,0009	0,0002	0,0016	0,0028	0,0015	0,0015	0,0004	0,0015	0,0013	0,002	0,0024	0,0002	0,0003	0,001	0	0,0011	0,0009	0,0003	0	0	0	0,0001	0,0012	0,000
ф	0,0003	0	0	0	0	0,0006	0	0	0,0004	0	0	0,0001	0	0	0,0003	0	0,0002	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	1 1
x	0,0008	0,0001	0,0005	0,0002	0,0002	0,0001	0	0,0001	0,0006	0	0,0003	0,0003	0,0002	0,0006	0,0024	0,0004	0,0004	0,0004	0,0001	0,0003	0	0	0	0,0001	0	0	0	0	0	0	0	
ц	0,0007	0	0,0001	0	0	0,001	0	0	0,0004	0	0	0	0	0	0,0004	0,0001	0	0	0	0,0002	0	0	0	0	0	0	0	0,0002	0	0		(
ч	0,0027	0	0,0001	0	0	0,0045	0	0	0,0018	0	0,0006	0	0	0,001	0,0001	0,0001	0,0001	0,0001	0,0037	0,0006	0	0	0	0	0,0001	0	0	0	0,0003	0	0	(
w	0,0012	0	0,0001	0	0	0,0024	0	0	0,0021	0	0,0005	0,0005	0	0,0005	0,0003	0,0001	0	0	0,0001	0,0005	0	0	0	0	0	0	0	0	0,0004	0	0	
щ	0,0005	0	0	0	0	0,0015	0	0	0,0012	0	0	0	0	0,0001	0	0	0	0	0	0,0002	0	0	0	0	0	0	0	0	0	0	0	
ъ	0	0	0	0	0	0,0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,000
ы	0,0001	0,0005	0,0014	0,0002	0,0003	0,0012	0,0001	0,0002	0,0004	0,002	0,0005	0,0024	0,0016	0,0006	0,0003	0,0006	0,0004	0,0012	0,0008	0,0002	0	0,0011	0	0,0003	0,0007	0	0	0	0	0,0001	0	0,000
b	0,0002	0,0003	0,0014	0,0003	0,0005	0,0012	0,0001	0,0004	0,0008	0	0,0017	0,0002	0,0007	0,0027	0,0009	0,001	0,0002	0,0018	0,0009	0,0002	0,0001	0,0001	0,0002	0,0005	0,0004	0	0	0	0	0,0001	0,0004	0,000
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0029	0	0	0	0	0	0	0	0	0	0	0	0	1 1
ю	0,0001	0,0003	0,0003	0,0001	0,0004	0	0,0001	0,0001	0,0002	0	0,0003	0,0001	0,0002	0,0003	0,0002	0,0003	0,0002	0,0005	0,0005	0	0,0001	0,0001	0	0,0002	0,0001	0,0006	0	0	0	0	0,0001	
Я	0,0003	0,0004	0,002	0,0004	0,001	0,0003	0,0003	0,0006	0,0009	0,0001	0,001	0,0011	0,0008	0,002	0,0006	0,0012	0,0004	0,0019	0,002	0,0003	0,0001	0,0003	0	0,0005	0,0001	0,0005	0	0	0	0,0001	0,0002	0,000

Обчислюємо ентропію символів за формулою:

$$H(Z) = -\sum_{i=1}^{n} p_i \log p_i.$$

Надлишковість:

$$R = 1 - \frac{H_{\infty}}{H_0}$$

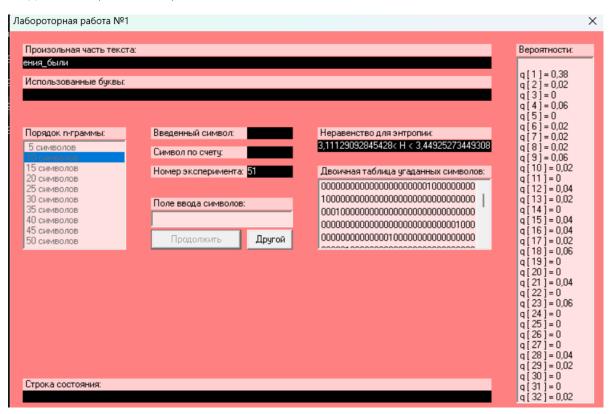
Метрика	3 пробілами	Без пробілів
H1	4,37434	4,45304
R1	0,13283	0,10939
Н2_Перетин	3,9973	4,14776
R2_Перетин	0,5862	0,57385
Н2_Без_Перетин	3,99743	4,14757
R2_Без_Перетин	0,58424	0,57197

### 2. За допомогою програми CoolPinkProgram оцінити значення

#### Порядок п-грами 10 символів

Ентропія 3,111 < H < 3,449

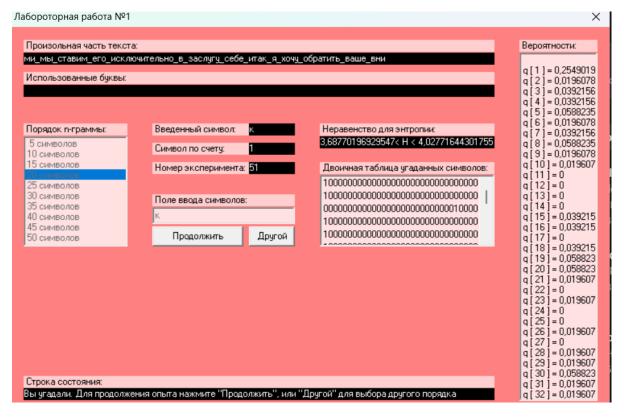
**Надлишок** 0,32 < R < 0,38



Порядок п-грами 20 символів

Ентропія 3,688 < Н < 4,028

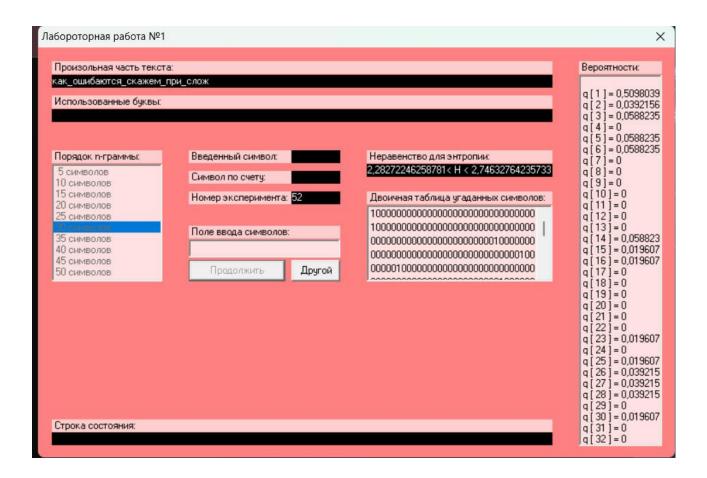
**Надлишок** 0,20 < R < 0,27



#### Порядок п-грами 30 символів

Ентропія 2,283 < H < 2,746

**Надлишок** 0,46 < R < 0,55



## Висновок

У ході роботи було проведено експериментальну оцінку ентропії та надлишковості джерела відкритого тексту російською мовою. Підрахунок частот символів і біграм показав природний нерівномірний розподіл, характерний для мови. Обчислені значення  $H_1$ ,  $H_2$  та ентропії для n-грам (n = 10, 20, 30) підтвердили наявність статистичних залежностей між символами: зі зростанням порядку n-грам ентропія зменшується, що вказує на передбачуваність тексту.

Таким чином, ми засвоїли методи оцінки ентропії та надлишковості текстів і отримали практичні навички їх експериментального визначення.