

# **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 б) частот біграм в тексті, а також підрахунку с)  $H_1$  та  $H_2$  за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення  $H_1$  та  $H_2$  на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення  $H_1$  та  $H_2$  на тому ж тексті, в якому вилучено всі пробіли.

А) Частота букв

*Частота букв рахується за формулою:*

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

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

```
[Running] python -u "d:\3 курс\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
г	0,016126
д	0,023597
е	0,068249
ж	0,007564
з	0,014923
и	0,057457
й	0,009841
к	0,030708
л	0,044157
м	0,025196
н	0,05383
о	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
з	0,017771
и	0,068421
й	0,011719
к	0,036568
л	0,052584
м	0,030004
н	0,064102
о	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

*b) частот біграм в тексті*

## Біграми які перетинаються з пробілами

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ь	э	ю	я			
а	0.0046	0.0028	0.0063	0.016	0.0044	0.0053	0.0028	0.0018	0.0046	0.0109	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	0.0021	0	0	0.0024	0	0.0016			
б	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	0.0007	0.0021		
в	0.0005	0.001	0	0	0	0	0	0	0	0.0007	0	0.0002	0.0006	0.0001	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	0.0003		
г	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	0.0001	0	0	0.0026	0.0005	0	0.0003		
д	0.0005	0.0021	0	0	0	0.0009	0.0005	0	0	0.0006	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		
е	0.0014	0.0042	0.0001	0.0009	0	0	0.0038	0.0001	0	0.0002	0	0.0003	0.0004	0.0001	0.0019	0.0033	0.0001	0.0001	0.0004	0.0002	0.0015	0	0	0.0002	0.0001	0.0002	0	0.0001	0.0005	0.0005	0	0.0004		
ж	0.0003	0.0003	0.0001	0.0001	0.0001	0.0013	0.0001	0.0013	0.0018	0	0.0002	0.0014	0.0001	0.0002	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	0.0002	0.0002			
з	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	0.0002	0	0	0	0	0	0	0	0	0	0	0	0		
и	0.0016	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	0.0001	0.0017	0.0005	0.0002	0	0	0.0002	0.0001			
й	0.0078	0	0	0	0	0	0	0	0	0	0	0	0	0.0004	0	0	0.0004	0.0004	0	0	0	0	0.0001	0.0002	0.0002	0	0	0	0	0	0	0		
к	0.0042	0.0071	0	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		
л	0.0078	0.0075	0.0001	0	0.0001	0.0001	0.0044	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	0	0.0001	0	0	0.0009	0.0035	0	0.0007	0.0013
м	0.0072	0.0036	0	0.0001	0.0001	0	0.0031	0	0.0028	0	0.0001	0.0002	0.0001	0.0013	0.0029	0.0002	0.0003	0.0001	0.0017	0	0	0	0	0	0	0	0	0	0.0007	0	0.0005	0.0005		
н	0.0033	0.0091	0	0	0	0.0001	0.0009	0.0088	0	0.0001	0.0007	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	0.0032	0.0009	0	0.0002	0.0014	
о	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	0.0005	0.0001	0.0003	0.0005	0.0001	0.0017	0.0008	0.0002	0	0	0.0001	0.0004	0.0005		
п	0	0.0017	0	0	0	0	0	0.0021	0	0	0.0013	0	0.0001	0.0009	0	0.0003	0.0009	0.0001	0.0003	0.0002	0.0001	0.0005	0	0	0	0	0	0	0	0.0003	0.0001	0	0.0004	
р	0.0014	0.0075	0.0001	0.0001	0.0001	0.0021	0.0056	0.0002	0	0.0007	0	0.0002	0.0005	0.0002	0.0001	0.0003	0.0002	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	0.0002	0.0001	
с	0.0029	0.0017	0.0001	0.0001	0.0001	0.0016	0.0009	0	0.0018	0	0.0039	0.0025	0.0006	0.0009	0.0026	0.0016	0.0002	0.0009	0.0092	0.0007	0.0007	0.0001	0.0002	0.0003	0.0001	0	0	0.0003	0.0004	0.0001	0.0007	0.0037		
т	0.0056	0.0056	0	0.002	0	0.0002	0.0057	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	0.0004	0	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	0.0002	0.0012	0.0003	0	0.0001	0.0001	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	0.0001	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	
х	0.0029	0.0006	0	0.0002	0.0001	0	0.0001	0	0	0.0003	0	0	0.0001	0.0001	0.0003	0.0018	0	0.0002	0	0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001		
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	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	0.0003	0	0	0
ш	0.0001	0.0001	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	0.0003	0	0	
щ	0.0001	0.0004	0	0	0	0	0.0012	0	0	0.0001	0	0	0	0	0	0.0001	0	0	0	0.0001	0.0001	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	0.0002	0	
ь	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	0.0008	0	0.0002	0.0005	0	0	0	0	0	0	0	0	
э	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	0.0007	0.0003	0	0	0.0001	0.0001	0.0003	0	0	0	0	0.0003	0.0003	0	0
ю	0.0025	0	0.0002	0	0	0	0.0002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0001	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	0.0001	0	0.0002	0	0.0004	0	0	0	0	0.0001	0.0001		

## Біграми які не перетинаються з пробілами

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ь	э	ю	я		
	0.0046	0.0028	0.0062	0.016	0.0045	0.0052	0.0028	0.0019	0.0045	0.0109	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	0.0024	0	0.0015		
а	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	0.0009	0.0008	0.0004	0	0	0.0008	0.0022		
б	0.0002	0.001	0	0	0	0	0.0023	0	0	0.0007	0	0.0002	0.0006	0.0001	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	0.0003	
в	0.0068	0.0063	0	0.0001	0	0.0002	0.0056	0	0.0004	0.0003	0	0.0002	0.0006	0.0001	0.0009	0.0063	0.0002	0.0006	0.0021	0.0003	0.0007	0	0.0001	0.0001	0.0001	0.0001	0	0	0.0025	0.0005	0	0.0004	
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	0.0005	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	0.0002	0.0001	0.0002	0	0.0001	0.0005	0.0005	0	0	0.0004	
е	0.0158	0.0001	0.0009	0.0022	0.003	0.0029	0.0012	0.0007	0.0013	0.0003	0.0018	0.0015	0.0059	0.0038	0.0063	0.0002	0.0012	0.0065	0.0041	0.0052	0.0002	0.0001	0.0004	0.0001	0.0002	0.0006	0	0	0.0002	0.0004	0	0.0003	
ж	0.0013	0.0013	0	0	0	0.0001	0.0023	0	0	0.0011	0	0.0002	0.0001	0.0001	0.0001	0.0001	0	0.0001	0	0.0001	0	0	0	0	0	0	0	0	0	0	0	0.0002	0.0004
з	0.0016	0.0056	0.0002	0.0011	0.0003	0.0009	0.0007	0.0001	0.0001	0.0005	0	0.0011	0.0004	0.0002	0.0014	0.0007	0.0001	0.0003	0	0	0.0005	0	0	0	0	0	0	0	0	0.0004	0.0001	0	0.0003
и	0.0158	0.0002	0.0005	0.0009	0.0005	0.0015	0.0017	0.0002	0.0021	0.0005	0.0016	0.0026	0.0062	0.0025	0.0032	0.0004	0.0003	0.0011	0.0029	0.0041	0.0001	0.0001	0.0014	0.0001	0.0017	0.0005	0.0002	0	0	0.0002	0.0003	0.0004	
й	0.0077	0	0	0	0	0.0001	0	0	0	0	0.0002	0	0	0.0001	0.0004	0	0	0.0004	0.0004	0	0	0	0	0	0	0	0	0	0	0	0	0	0
к	0.0041	0.0006	0.0008	0.0008	0.0001	0.0013	0.0009	0.0001	0.0009	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	
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	0.0068	0.0001	0	0.0019	0.0002	0.0012	0	0	0	0.0001	0	0	0	0.0008	0.0034	0	0.0007	0.0001
м	0.0071	0.0036	0	0.0001	0.0001	0	0.0051	0	0	0.0029	0	0.0001	0.0002	0.0001	0.0013	0.0023	0.0002	0.0001	0.0002	0.0001	0.0017	0	0	0	0	0	0	0	0	0.0007	0	0	0.0003
н	0.0032	0.0009	0	0	0.0001	0.0009	0.0088	0	0	0.0071	0	0.0003	0	0	0.0031	0.0097	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	0.0004	
о	0.0203	0.0013	0.0023	0.0059	0.004	0.0013	0.0017	0.0001	0.0012	0.0004	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	
п	0.0017	0	0	0	0	0.0021	0	0	0.0012	0	0.0001	0.0009	0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	
р	0.0015	0.0075	0.0001	0.0005	0.0001	0.0002	0.0057	0.0002	0	0.0057	0	0.0004	0.0003	0.0002	0.001	0.0009	0.0001	0.0001	0.0009	0.0009	0.0024	0	0.0002	0.0001	0.0002	0.0004	0	0	0.0013	0.0006	0.0002	0.0003	
с	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	0.0024	0.0017	0.0002	0.0009	0.0092	0.0007	0	0.0002	0.0001	0.0003	0.0004	0	0	0.0003	0.0004	0	0.0001	
т	0.0017	0.0007	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	
у	0.0063	0.0001	0.0007	0.0007	0.0009	0.0019	0.0002	0.0012	0.0003	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	
ф	0.0003	0	0.0003	0	0	0	0.0005	0	0	0.0003	0	0.0003	0	0.0001	0	0.0003	0	0.0002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
х	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	0.0001	0	0.0001	0.0001	0	0.0001	0.0001	0	0.0001	0.0001	0	0.0001	0	0	0	0	0	0	0	0	0	0	0	0
ч	0.0006	0.0002	0.0001	0	0	0.0038	0	0	0.0016	0	0.0005	0	0	0.0008	0.0001	0	0	0.0003	0.0005	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ш	0.0001	0.0001	0.0001	0	0.0001	0	0.0021	0.0001	0.0018	0	0.0001	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	0	0	0
щ	0	0.0004	0	0	0	0.0012	0	0	0.0001	0	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	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
ь	0.0038	0	0.0003	0.0008	0.0001	0.0001	0.0009	0.0001	0.0001	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	0
э	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	0.0001	0.0001	0	0	0	0.0001	0.0001	0.0003	0	0	0	0	0	0	0	0
ю	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	0	0	0	0	0	0	0
я	0.0021	0	0.0001	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	0	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	0.0002	0	0.0002	0	0.0002	0	0.0004	0	0	0	0	0

*Надлишковість:*

Метрика	З пробілами	Без пробілів
H1	4,37434	4,45304
R1	0,13283	0,10939
H2_Перетин	3,9973	4,14776
R2_Перетин	0,5862	0,57385
H2_Без_Перетин	3,99743	4,14757
R2_Без_Перетин	0,58424	0,57197

## Порядок n-грами 10 символів

**Ентропія**  $3,111 < H < 3,449$

**Надлишок**  $0,32 < R < 0,38$

## Порядок n-грами 20 символів

**Ентропія**  $3,688 < H < 4,028$

**Надлишок**  $0,20 < R < 0,27$

[illegible]

### Порядок n-грами 30 символів

**Ентропія  $2,283 < H < 2,746$**

**Надлишок**  $0,46 < R < 0,55$

[illegible]

## **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.