



Комп'ютерний практикум №1
Експериментальна оцінка ентропії на символ
джерела відкритого тексту

Роботу виконали:

Біла Анастасія і Лета Яна,
студенти 3 курсу ФТІ НТУУ «КПІ»,
спеціальність «Кібербезпека», група ФБ-02

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.

Постановка задачі:

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку 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.
2. За допомогою програми CoolPinkProgram оцінити значення $H^{(10)}$, $H^{(20)}$, $H^{(30)}$.
3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Хід роботи:

Обираємо текст «Проблеми розвитку психіки», автор Леонтьєв Олексій, текст написаний російською мовою.

Текст містить багато символів окрім власне літер, тому проводимо попередню фільтрацію тексту та створюємо два файли: один містить пробіли, інший - ні.

```
def check(text): # функція очищення тексту для подальшого використання у завданні
    reg = re.compile('^[^а-яёА-ЯЁ ]+') # забирає спецсимволи і навіть цифри
    new = reg.sub('', text)
    return new

splited = check(text).split(' ')
arr = [word.rstrip('.').lower() for word in splited if word]
cleansp = ' '.join(arr)
clean = ''.join(arr)
clean_textsp = open('clean_textsp.txt', 'w')
clean_textsp.write(cleansp) # текст з пробілами
clean_text = open('clean_text.txt', 'w')
clean_text.write(clean) # текст без пробілів
```

Створюємо функцію для підрахунку частоти букв у тексті, а також обчислення ентропії та надлишку. Результати частоти букв експортуємо у файл «frequency_of_letters.xlsx».

```

def letters_frequency(txt): # функція для підрахунку частоти букв
    c = Counter()
    frequency = {}
    total = 0
    for line in txt:
        c += Counter(line)
    for letter in alphabet_without_spaces:
        frequency[letter] = c[letter]/len(txt)
        #print(letter, frequency[letter])
    # h1 обчислюємо ентропію
    for i in frequency.values():
        total += i*math.log2(i)
    h1 = -total
    #print('H1:', h1)

    # обчислюємо надлишок
    R = 1 - (h1 / math.log2(len(alphabet_without_spaces)))
    #print('R:', R)

    letters_data = pd.DataFrame(data=frequency, index=['частота'])
    print(letters_data)
    letters_data.to_excel('frequency_of_letters.xlsx')

```

Для тексту з пробілами:

Частота букв (відсортована за спаданням частот):

A	В	С	Д	Е	F	G	Н	I	Ј	К	L	М	Н	О	Р	Q	R	S
о	е	и	н	т	а	с	в	р	л	м	к	п	а	я	у	ы	з	
частота	0,097366	0,084177	0,070685	0,060334	0,059946	0,053974	0,050193	0,042156	0,038166	0,029304	0,026881	0,025419	0,023973	0,022788	0,019951	0,018921	0,017591	0,01539

S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH			
з	ч	ь	г	б	й	х	ж	ю	щ	ш	ц	э	ф	ъ	ë			
0,01539	0,01331	0,01326	0,01239	0,01153	0,01145	0,01025	0,00808	0,00649	0,00539	0,00458	0,00442	0,00414	0,00313	0,00066	2E-05			

Ентропія:

H1: 4.035959054535917

Надлишковість:

R: 0.19991202926681495

Для тексту без пробілів:

Частота букв (відсортована за спаданням частот):

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
о	е	и	н	т	а	с	в	р		л	м	к	п	д	я	у	ы	з
частота	0,11239	0,097166	0,081592	0,069643	0,069195	0,062303	0,057938	0,048661	0,044055	0,033826	0,031029	0,029342	0,027672	0,026304	0,023029	0,02184	0,020305	0,017764

S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH		
з	ч	б	г	б	й	х	ж	ю	щ	ш	ц	э	ф	ъ	ë		
0,01776	0,01537	0,01531	0,01431	0,01331	0,01322	0,01183	0,00933	0,00749	0,00622	0,00528	0,0051	0,00478	0,00361	0,00076	2,3E-05		

Ентропія:

H1: 4.451702489104348

Надлишковість:

R: 0.11749510768391025

Створюємо функцію для підрахунку частоти біграм у тексті, а також обчислення ентропії та надлишку. Результати частоти біграм експортуємо у файл «frequency_of_bigrams.xlsx».

```
def bigrams_frequency(txt, intersection=True): #частота біграм
    c = Counter()
    frequency = {}
    index = 0
    if txt == clean:
        alph = alphabet_without_spaces
    else:
        alph = alphabet_with_spaces
    for letter1 in alph:
        for letter2 in alph:
            bigram = letter1 + letter2
            c[bigram] = 0

    if intersection is True: # біграми з перетином
        for i in range(len(txt) - 1): # до останньої літери
            bigram = txt[i] + txt[i + 1]
            c[bigram] += 1 # рахуємо скільки разів зустрічається біграма
        for bigram in c.keys():
            frequency[bigram] = c[bigram] / (2 * sum(c.values())) # частота кожної біграми
            #print(bigram, frequency[bigram])

    else: # біграми без перетину
        for i in range(0, len(txt) - 2, 2): # крок 2
            bigram = txt[i] + txt[i + 1]
            c[bigram] += 1 # рахуємо скільки разів зустрічається біграма

        for bigram in c.keys():
            frequency[bigram] = c[bigram] / (2 * sum(c.values())) # частота кожної біграми
            #print(bigram, frequency[bigram])
```

```

# h2 обчислюємо ентропію
total = 0
for i in frequency.values():
    if i > 0:
        total += i*math.log2(i)
h2 = -total
#print('H2:', h2)

# обчислюємо надлишок
R = 1 - (h2 / math.log2(len(alph)))
#print('R:', R)
big_fr = list(frequency.values())
bigrams_data = pd.DataFrame(data=frequency, index=al, columns=al)

for i in range(0, len(al)):
    bigrams_data[al[i]] = big_fr[index:len(al) + index]
    index = len(al) + index
print(bigrams_data)

bigrams_data.to_excel('frequency_of_bigrams.xlsx')

```

Для тексту з пробілами:

Частота біграм з перетином:

	а	б	в	г	д	е	ё	з	ж	з	и	ы	й	к	л	м
а	0,00053	0,00076	1,9E-06	0	4,8E-07	4,4E-05	0,00055	0	1,4E-06	0	0,00043	0,00241	0,00047	8,1E-06	1,9E-06	
б	0,0008	0,0044	0,0007	3,3E-06	2,6E-05	0	0	0,0003	0	0	0	0	8,1E-06	0,00085	0,0009	0,00045
в	0,0013	1,4E-06	0,00012	6,8E-05	0	0	1,5E-05	0,00027	0	0	3,8E-06	0	0,00015	9,6E-07	6,3E-05	0,00033
г	0,00656	0,00033	9,9E-05	0,00341	0,00035	1,9E-06	0,00014	0	0	0	9,6E-07	2,9E-06	0,00191	1,4E-06	4,8E-07	0,00103
д	0,00062	0,00125	1,4E-06	0,00255	0,00012	0	3,8E-06	0	0	0	0	1,4E-06	0,00015	0,0014	9,6E-07	0,00033
е	0,00331	0,0002	3,8E-06	2,9E-06	0,00056	0,00037	0,00012	7,2E-05	0	0	0	1,1E-05	0,00157	4,8E-07	9,6E-07	0
ё	0,00136	0,00064	4,8E-07	2,3E-05	0	0,00159	0,00829	0,00045	0	0	0	0	0,00018	9,3E-05	2,4E-06	4,8E-07
з	0	0,00146	2,4E-06	1,4E-06	3,3E-06	8,1E-06	0,00024	2,4E-06	0	0	1,4E-06	1,9E-06	4,6E-05	0,00012	3,8E-05	0,00032
ж	0,00188	0	0,00077	2,9E-05	4,3E-06	0,00038	0,00065	0	6,7E-06	0	0	8,2E-05	0	0	0,00014	9,6E-07
з	0,00081	5,7E-06	0	0,00258	0,00025	2,9E-05	0,00033	9,6E-07	0	6,2E-06	0	4,3E-06	0	0	4,8E-07	8,6E-06
и	0,0015	0,00071	0	4,8E-07	0,00018	2,6E-05	0,00115	0	4,8E-07	0,00048	0,00042	5,1E-05	0,00042	0	0	0
ы	0,00514	0,00192	1,9E-06	4,8E-06	0	0,0031	0,00227	0	2,6E-05	4,8E-07	0,00153	0,00751	0,00302	0	0	0
й	0	0,00017	2,4E-06	0	0	0	0,00082	0	9,6E-07	0	4,4E-05	0,00033	0,00202	1,8E-05	0	0
к	3,8E-06	0	0,00017	0,00017	0	9,6E-07	4,8E-07	0	1,9E-06	0	0,00095	0,00022	4,8E-07	0,00436	1,9E-06	4,8E-07
л	0,00333	9,9E-05	0,00083	0,00251	4,8E-07	3,8E-06	7,2E-06	0	0	0,00046	9,7E-05	0,00165	6,2E-05	4,8E-07	0,00188	0
м	0,00078	0,00255	0	0,00128	0,00076	1,9E-06	0,0004	0	0	0,00146	0,00061	0,00043	0,00048	1,4E-06	0,00248	0,0004
н	0,00215	0,00181	3,9E-05	0	0	0,00104	0,0005	0	0	0	9,6E-05	0,00046	2,4E-05	6,2E-06	0	0,00097
о	0,00471	0,00105	0,00042	0,00011	0	0,00018	6,6E-05	0	0	0	0	0,00249	7,2E-05	0	6,7E-05	9,6E-07
п	0,0063	0,00303	1,9E-05	0,0012	2,3E-05	0	0	0	0	0	0	0	0,00099	2,4E-05	9,6E-07	9,6E-07
р	0,00738	5,6E-05	0,00027	0,00011	0,00024	6,9E-05	0,00182	0	0	0	7,6E-06	1,2E-05	0	6,2E-06	1,9E-06	8,6E-06
с	0,00277	0,00063	0,00105	0,00118	3,8E-06	0,00046	0,0011	0	0	0,00094	7,2E-06	0,00022	1,4E-06	0	0,00014	3,3E-06
т	0,00727	0,00067	4,8E-07	0,00356	8,4E-05	0,00019	0,00293	0	2,9E-06	0	0,00032	0,00177	1,9E-06	6,7E-06	3,8E-06	0,00263
у	0,00307	0,00089	0,00058	0,00012	0,00336	0,00065	0,00236	4,3E-06	0,00019	0	0,00031	0,00115	7,7E-05	0	4,8E-07	9,6E-07
ф	0,00167	0,00177	0,00013	0,00023	9,6E-07	0,00111	0,00663	1,4E-06	8,4E-05	1,9E-05	0	0	4,3E-06	9,6E-07	0,00013	1,9E-06
х	0,00092	9,9E-05	9,6E-07	0,00082	0,00046	5,2E-05	0,00031	4,8E-07	1,3E-05	0	5,4E-05	0,00096	0	5,7E-06	7,2E-06	0,00012
ц	0,00034	4,5E-05	0,00034	0,0002	9,6E-07	0,00059	0,00049	0	6E-05	0	0,00019	0,00078	0,00057	0	0,00138	0
ч	0,00021	0,00029	0	0,00051	0	0,00054	0,00269	0	0	0,00062	0,0006	0,00073	8,9E-05	0	0	0,00213
ш	0,00195	0,00033	0,00011	4,8E-07	0,00014	2,9E-05	0,00326	4,8E-07	6,2E-06	3,3E-06	0,00123	0,00223	0,00027	3E-05	0	0,00011
щ	0,00016	0,00073	4,8E-07	2,1E-05	0	0,00056	0,00347	0	2,9E-06	4,8E-07	0,00037	0,00151	0,00089	3,3E-06	8,1E-06	0
ъ	6,2E-06	0,00019	0	1,4E-06	0	0	2,3E-05	0	1,9E-06	0	4,8E-07	0,00032	2,4E-05	1,2E-05	0,00031	6E-05
ь	0	0,00016	1,4E-06	3,8E-06	0	3,7E-05	9,3E-05	0	0,00161	4,8E-07	0,00025	0,00023	4,3E-06	8,5E-05	9,6E-07	0,0001
ю	9,6E-07	0	0,00041	9,7E-05	4,3E-06	7,2E-06	0,0003	0	0	0	3,8E-06	0,00072	8,8E-05	8,6E-06	0,00013	4,8E-07
я	1,3E-05	0	0,00031	1,4E-06	4,8E-07	7E-05	0,00025	0	6E-05	5E-05	2,4E-06	0,00173	0,00018	1,4E-05	0,0037	0,00017

н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ь	ю	я	
а	0,00246	0,00021	0	0,00232	5,3E-06	0	0,00015	0,0035	3,9E-05	0	9,6E-07	0	0,00285	0	0	0,23E-05	
б	2,9E-06	0,00186	0,00236	0,00232	0,00021	0,00012	0	0,00039	7,6E-06	0	0	4,8E-07	4,8E-07	0,00096	0	0,86E-06	
в	4,8E-07	7,5E-05	0,0073	0,00162	0	5,3E-06	0,00157	0	0	0,00028	0	0	0	9,6E-07	0,00147	0	0,62E-06
г	6,5E-05	4,8E-06	9,6E-07	0,0003	4,7E-05	0,00047	0,0019	0,00018	2,4E-06	4,8E-07	0,00026	0	0	0	4,8E-07	0,00025	0,53E-06
д	3,3E-06	0,00011	6,7E-06	0,00123	0,00352	0,00026	0,00033	0,0001	0,00037	0	0	0,00119	0	0	0	0,78E-05	
е	0,00052	1,1E-05	0,00022	0,00298	7E-05	0,00338	0,00048	5,7E-05	0,00026	4,8E-07	0	2,4E-05	0,00082	0	0	0	0,96E-07
ё	4,8E-07	0,00074	0,00069	0,00402	0,00428	1,8E-05	0,00231	0,00118	0,00018	8,7E-05	1,4E-06	0	0	0,00045	0	0	0,96E-07
э	0	1,9E-06	0,00057	0,00378	0,00048	1E-05	0,00118	0,00587	0,00024	6,7E-06	1,1E-05	1,4E-06	0	0	0,00075	0	0,48E-07
ж	0	1,4E-06	1,2E-05	9,6E-06	6,1E-05	4,6E-05	0,00045	1,4E-05	4,8E-07	1E-05	1,4E-06	0	0,00012	0	0	0	0,54E-05
з	4,8E-06	9,6E-07	0	6,2E-05	0,0002	0,00021	0,00102	0,00141	0,00036	0,00043	4,9E-05	0	1,1E-05	0,00019	0	0	0,18E-05
и	4,8E-07	2,4E-06	0,00021	0,0001	4,8E-07	0,00123	0,00737	0,00154	0,00021	0	0,00101	4,8E-07	0	6,7E-05	0	0	0
ы	9,6E-07	0	2,4E-05	0,00059	0	1,8E-05	0,0005	9,1E-06	0,00066	2,3E-05	3,8E-06	7,2E-06	0,00058	4,8E-07	0	0	0
й	0	4,8E-07	9,6E-06	0,00059	2,6E-05	4,2E-05	5,9E-05	0,00081	0,00049	2,3E-05	5,2E-05	0	7,6E-06	0,00018	0	0	0,00034
к	0,00324	0	8,1E-06	0,00055	6,2E-06	1,6E-05	0,00021	4,8E-07	2,9E-06	4,8E-07	4,1E-05	4,8E-07	1,4E-06	4,5E-05	4E-05	0	0
л	0,00042	6,2E-06	0	0,00022	6,2E-06	1,1E-05	5,3E-06	4,3E-06	7,6E-06	0,0002	1,1E-05	0	1,6E-05	0	0	0	0,72E-05
м	0,00123	3,8E-06	0,00022	0	0	0,00014	7,8E-05	1,9E-06	0,0001	0,00012	1,3E-05	0	4,8E-07	0	0	0	0,00231
н	0,00042	0,00014	4,4E-05	0	0	1,4E-06	9,2E-05	4,9E-05	5,7E-06	0	6,2E-06	1,5E-05	0,00094	0	0	0	1,4E-06
о	0,00088	0,00054	0,00042	4,7E-05	3,8E-06	0	0	9,6E-07	0,00069	0	4,8E-07	0	0,00026	1,5E-05	0	0	0,48E-06
п	7,2E-06	0,00424	0,0084	0,00026	0	0,00012	4,8E-06	4,8E-06	6,9E-05	0	0	0	0	2,9E-05	0	0	0,48E-07
р	1,1E-05	0	9,1E-06	0,00012	2,6E-05	6,7E-06	0,00042	0	0,00055	0	9,6E-07	0	0	0	0,0002	0	0,0003
с	6,2E-06	4,8E-06	0,00286	0,0005	0,00045	0,00027	2,5E-05	0,00213	0	0	0	0	0	0	0	0	0,00019
т	3,3E-06	3,6E-05	0,00444	3,3E-06	0,0042	0,00126	0,00195	9,6E-07	0	0	0	0	0	0	2,4E-06	0	0
у	0,00226	0,00013	0,0028	4,8E-07	1,2E-05	0,00069	0,00348	0,00013	0,00093	0	0	0	0	1E-05	0	0	0,14E-06
ф	0	0,00322	0,00216	0	0,00035	2,5E-05	0,00283	0,00183	2,6E-05	0	0	0	0	1,9E-06	0	0	0,48E-07
х	4,8E-06	4,8E-07	0,00121	4,8E-07	0,00033	0,00102	1,2E-05	0,00015	2,8E-05	0	4,8E-07	0	0	0	0	0	0,96E-06
ц	0	1,4E-06	4,8E-07	0,00143	1,3E-05	3,3E-06	0,00278	0,00017	0,00019	0,00331	4,8E-07	0	4,3E-05	0	0	0	0,43E-06
ч	1,9E-06	2,4E-06	0,00013	0	0,00399	5,6E-05	4,3E-06	0,00026	0	0,00029	7E-05	0	0	0,00016	0	0	0,00015
ш	0,00158	3,8E-06	0,00098	4,8E-07	0	0,00086	0,00011	0,00043	0	4,8E-07	7,6E-05	4,3E-05	0	0	3,9E-05	0	0,96E-07
щ	0,00123	0,00643	0,00125	0	4,8E-07	9,6E-07	0,00336	0,00044	0	3,1E-05	0	0,00095	1,1E-05	0	0	0	0
ъ	4,8E-07	0,00283	0,00051	4,8E-07	0,0002	4,8E-07	4,8E-07	0,00055	0	9,6E-07	1,3E-05	0	0,00017	4,3E-06	0	0	0
ь	1,7E-05	4,8E-07	0	0,00031	9,6E-07	1,9E-06	2,4E-06	0	0,00017	3,8E-06	0	3,3E-06	0	0,00019	0	0,76E-05	0,0002
ю	5,3E-05	0,00062	0,00227	0,00067	0,00239	9,6E-07	0	3,8E-06	0	2,3E-05	0	0	4,8E-07	0	0	0,00278	0,71E-05
я	2,5E-05	0	0,00065	0	0,00079	0,00108	3,8E-06	0,00044	0	0	0,00081	4,8E-07	4,8E-07	9,6E-07	0	0	0,00017

H2: 4.476128294272646

R: 0.12016491639424842

Частота біграм без перетину:

а	б	в	г	д	е	ё	э	ж	з	и	ы	й	к	л	м	
а	0	0,00055	0,00075	3,8E-06	0	0	3,9E-05	0,00054	0	1,9E-06	0	0,00043	0,00241	0,00048	6,7E-06	2,9E-06
б	0,00082	0,0044	0,00069	1,9E-06	3,2E-05	0	0	0,00028	0	0	0	9,6E-06	0,00085	0,0009	0,00044	
в	0,00134	1,9E-06	0,00011	6E-05	0	0	1,6E-05	0,00025	0	0	0	2,9E-06	0	0,00016	1,9E-06	7,8E-05
г	0,00652	0,00037	8,6E-05	0,0034	0,00036	3,8E-06	0,00015	0	0	0	0	2,9E-06	0,00191	1,9E-06	0	0,00108
д	0,0006	0,00124	1,9E-06	0,00256	0,00012	0	3,8E-06	0	0	0	0	0	0,00014	0,00139	9,6E-07	0,00032
е	0,00326	0,00019	2,9E-06	3,8E-06	0,00056	0,00037	0,00013	6,7E-05	0	0	0	1,1E-05	0,00154	9,6E-07	9,6E-07	0
ё	0,00136	0,00062	9,6E-07	2,7E-05	0	0,00165	0,00832	0,00042	0	0	0	0	0,0002	9,5E-05	1,9E-06	9,6E-07
э	0	0,00147	1,9E-06	1,9E-06	3,8E-06	5,7E-06	0,00023	2,9E-06	0	0	0	2,9E-06	3,8E-06	4,3E-05	0,00012	3,4E-05
ж	0,00188	0	0,00077	3,2E-05	4,8E-06	0,00038	0,00062	0	8,6E-06	0	0	9,1E-05	0	0	0,00015	9,6E-07
з	0,0008	3,8E-06	0	0,00258	0,00025	3E-05	0,00034	9,6E-07	0	5,7E-06	0	3,8E-06	0	0	0	4,8E-06
и	0,00153	0,0007	0	9,6E-07	0,00018	3E-05	0,0012	0	0	0,00047	0,00041	4,9E-05	0,00042	0	0	0
ы	0,00518	0,00186	2,9E-06	2,9E-06	0	0,00308	0,00233	0	2,1E-05	0	0,00152	0,00741	0,00301	0	0	0
й	0	0,00016	0	0	0	0	0,00085	0	0	0	4,3E-05	0,00031	0,00197	1,7E-05	0	0
к	4,8E-06	0	0,00017	0,00015	0	0	0	0	1,9E-06	0	0,00094	0,00022	0	0,00436	1,9E-06	9,6E-07
л	0,00328	9,9E-05	0,00084	0,00252	9,6E-07	2,9E-06	7,6E-06	0	0	0,00045	9,1E-05	0,00164	5,2E-05	0	0,00187	0
м	0,00081	0,00258	0	0,0013	0,00078	2,9E-06	0,0004	0	0	0,00146	0,00063	0,00041	0,00048	1,9E-06	0,00252	0,00042
н	0,00216	0,00178	3,5E-05	0	0	0,00099	0,0005	0	0	0	0	9,4E-05	0,00046	1,9E-05	5,7E-06	0
о	0,0047	0,00106	0,00042	0,00013	0	0,00019	7,3E-05	0	0	0	0	0,00247	6,5E-05	0	6,3E-05	9,6E-07
п	0,00621	0,00301	1,4E-05	0,00119	2,8E-05	0	0	0	0	0	0	0	0,00097	2,6E-05	1,9E-06	1,9E-06
р	0,00742	5,8E-05	0,00028	9,8E-05	0,00025	6,6E-05	0,00179	0	0	0	0	6,7E-06	1,5E-05	0	7,6E-06	1,9E-06
с	0,00275	0,00063	0,00098	0,00112	3,8E-06	0,00046	0,00106	0	0	0,00098	9,6E-06	0,00023	9,6E-07	0	0,00014	9,6E-07
т	0,00734	0,00064	9,6E-07	0,00357	8,9E-05	0,00019	0,00296	0	4,8E-06	0	0,00032	0,00175	2,9E-06	8,6E-06	2,9E-06	0,00259
у	0,00309	0,00089	0,00057	0,00012	0,00332	0,00062	0,00235	4,8E-06	0,00019	0	0,00034	0,00113	8,5E-05	0	9,6E-07	9,6E-07
ф	0,00166	0,00174	0,00013	0,00023	0	0,0011	0,00662	9,6E-07	8,9E-05	1,8E-05	0	0	2,9E-06	9,6E-07	0,00012	9,6E-07
х	0,00092	9,8E-05	9,6E-07	0,00085	0,00047	4,5E-05	0,00031	0	1,4E-05	0	6,2E-05	0,0				

	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ѣ	ю	я
а	0,00252	0,00022	0	0,00234	4,8E-06	0	0,00017	0,00351	3,4E-05	0	9,6E-07	0	0,0029	0	0	0	2E-05
б	3,8E-06	0,00183	0,00232	0,00229	0,00021	0,00011	0	0,00039	8,6E-06	0	0	9,6E-07	0	0,00096	0	0	9,6E-06
в	9,6E-07	7,6E-05	0,00729	0,00163	0	4,8E-06	0,0016	0	0	0,00027	0	0	0	1,9E-06	0,00152	0	8,6E-06
г	6,6E-05	5,7E-06	9,6E-07	0,00031	4,7E-05	0,00047	0,00192	0,00017	1,9E-06	0	0,00025	0	0	0	0	0	3,8E-06
д	2,9E-06	9,9E-05	2,9E-06	0,00125	0,00349	0,00024	0,00032	9,8E-05	0,00036	0	0	0,00121	0	0	0	0	7,1E-05
е	0,00049	1,7E-05	0,00021	0,00297	7,5E-05	0,00334	0,00047	6,5E-05	0,00026	0	0	2,5E-05	0,00082	0	0	0	0
ѣ	9,6E-07	0,00075	0,0007	0,00408	0,00431	2E-05	0,0023	0,0012	0,00018	7,1E-05	9,6E-07	0	0	0,00044	0	0	0
э	0	1,9E-06	0,0006	0,00377	0,00048	8,6E-06	0,00116	0,00591	0,00025	8,6E-06	1,3E-05	2,9E-06	0	0	0,00073	0	9,6E-07
ж	0	9,6E-07	1,1E-05	1,1E-05	6,3E-05	4,6E-05	0,00044	1,2E-05	9,6E-07	7,6E-06	0	0	0,00011	0	0	0	5,3E-05
з	6,7E-06	9,6E-07	0	6,7E-05	0,0002	0,00022	0,00097	0,00138	0,00034	0,00044	5E-05	0	9,6E-06	0,00017	0	0	1,2E-05
и	0	1,9E-06	0,00021	0,00011	9,6E-07	0,00125	0,00729	0,00154	0,00021	0	0,00098	9,6E-07	0	7,2E-05	0	0	0
ы	9,6E-07	0	2,1E-05	0,00061	0	1,5E-05	0,00051	3,8E-06	0,00064	2,4E-05	2,9E-06	4,8E-06	0,00058	0	0	0	0
й	0	0	1,1E-05	0,00058	2,8E-05	4,3E-05	6,1E-05	0,00085	0,00046	2,1E-05	4,9E-05	0	7,6E-06	0,0002	0	0	0,00034
к	0,00321	0	8,6E-06	0,00057	2,9E-06	1,8E-05	0,0002	9,6E-07	1,9E-06	0	3,5E-05	0	1,9E-06	4,1E-05	3,6E-05	0	0
л	0,00042	4,8E-06	0	0,00023	5,7E-06	8,6E-06	4,8E-06	4,8E-06	7,6E-06	0,00019	1,3E-05	0	1,7E-05	0	0	0	7,5E-05
м	0,0012	4,8E-06	0,00025	0	0	0,00015	8E-05	9,6E-07	0,00011	0,00012	1,1E-05	0	9,6E-07	0	0	0	0,00235
н	0,00424	0,00015	4,9E-05	0	0	9,6E-07	9,4E-05	6,1E-05	5,7E-06	0	5,7E-06	1,6E-05	0,00093	0	0	0	0
о	0,00089	0,00053	0,00042	4,9E-05	4,8E-06	0	0	9,6E-07	0,00072	0	0	0	0,00025	1,5E-05	0	0	3,8E-06
п	6,7E-06	0,00433	0,00843	0,00026	0	0,00012	3,8E-06	3,8E-06	6,8E-05	0	0	0	0	2,6E-05	0	0	0
р	9,6E-06	0	7,6E-06	0,00013	3,3E-05	6,7E-06	0,00041	0	0,00055	0	1,9E-06	0	0	0	0,0002	0	0,0003
с	2,9E-06	5,7E-06	0,00293	0,00054	0,00044	0,00028	3,1E-05	0,00214	0	0	0	0	0	0	0	0	0,00019
т	2,9E-06	3,9E-05	0,0045	1,9E-06	0,00426	0,00125	0,00196	1,9E-06	0	0	0	0	0	0	3,8E-06	0	0
у	0,00223	0,00012	0,00277	9,6E-07	1,2E-05	0,0007	0,00348	0,00014	0,00092	0	0	0	0	1,2E-05	0	0	0,19E-06
ф	0	0,00322	0,0022	0	0,00032	2,9E-05	0,0028	0,00184	2,5E-05	0	0	0	0	1,9E-06	0	0	0
х	3,8E-06	0	0,00119	9,6E-07	0,00034	0,00101	1,1E-05	0,00014	2,6E-05	0	9,6E-07	0	0	0	0	0	7,6E-06
ц	0	1,9E-06	9,6E-07	0,00141	1,8E-05	5,7E-06	0,00278	0,00018	0,0002	0,00333	0	0	3,8E-05	0	0	0	5,7E-06
ч	2,9E-06	1,9E-06	0,00013	0	0,00409	5,7E-05	6,7E-06	0,00024	0	0,00027	7,4E-05	0	0	0,00015	0	0	0,00015
ш	0,00159	2,9E-06	0,00099	9,6E-07	0	0,00084	0,00011	0,00041	0	0	8,4E-05	4,1E-05	0	0	4,1E-05	0	0
щ	0,00124	0,00642	0,00124	0	0	1,9E-06	0,00335	0,00044	0	3,4E-05	0	0,00092	9,6E-06	0	0	0	0
ъ	0	0,00282	0,00052	0	0,0002	9,6E-07	9,6E-07	0,00053	0	0	1,1E-05	0	0,00017	6,7E-06	0	0	0
ь	1,5E-05	9,6E-07	0	0,00031	0	0	2,9E-06	0	0,00016	4,8E-06	0	3,8E-06	0	0,00019	0	7,1E-05	0,00022
ю	5,3E-05	0,00061	0,00227	0,00067	0,00238	9,6E-07	0	4,8E-06	0	2,8E-05	0	0	0	0	0	0,00275	6,6E-05
я	1,9E-05	0	0,00063	0	0,00079	0,00108	3,8E-06	0,00044	0	0	0,00079	0	9,6E-07	9,6E-07	0	0	0,00114

H2: 4.475432623719718

R: 0.12030165853362829

Для тексту без пробілів:

Частота біграм з перетином:

а	б	в	г	д	е	ѣ	ә	ж	з	и	ы	ѣ	к	л	м	
7,4E-05	0,00012	0,00298	0,00065	0,00184	0,00035	0	0	0,00055	0,00177	0,00049	3,3E-05	9,2E-05	0,00288	0,00113	0,00107	
б	0,00048	2,2E-06	6,6E-05	5,5E-06	1,8E-05	0,00094	2,8E-06	1,1E-06	1,1E-06	6E-05	0,00042	0,00014	6E-05	0,00011	8,8E-06	8,7E-05
в	0,00198	1,1E-05	0,00021	1,8E-05	0,00049	0,00143	5,5E-07	3E-05	5,5E-07	0,00113	0,00275	0,00081	0,00037	0,00025	4,7E-05	0,00046
г	0,00028	5E-06	3E-05	6,6E-06	3,7E-05	0,00144	0	3,3E-06	5,5E-07	0,00012	0,00057	4,9E-05	6,8E-05	1,3E-05	1,8E-05	4,7E-05
д	0,00098	1E-05	0,00025	0,0003	4,4E-05	0,00307	0	2,8E-06	0,00053	0,00073	0,00094	0,0002	0,00035	0,00011	2,3E-05	0,00023
е	0,00177	0,00089	0,00308	0,0002	0,00358	0,00109	0	0	0,00168	0,00012	0,00305	0,00118	8,7E-05	0,00019	0,00305	0,00269
ѣ	0	0	5,5E-07	0	0	5,5E-07	0	0	0	0	0	0	0	0	4,4E-06	1,1E-06
ә	0,0002	3E-05	0,00019	1,7E-06	1E-05	0,00026	0	0	0	3,9E-05	0,00028	5,1E-05	0,00012	6,1E-05	1,5E-05	0,00012
ж	0,00086	2,2E-06	2E-05	1,1E-06	7,7E-06	0,00057	0	0	0	1,3E-05	0,00035	2,3E-05	9,4E-05	0,00019	0,00014	5,4E-05
з	0,00235	3,9E-06	0,00028	1,1E-06	1,5E-05	0,00085	0	0	0	1,4E-05	0,00227	0,00013	9,4E-05	3,8E-05	9,9E-06	9,5E-05
и	0,00064	0,00021	0,00314	0,00088	0,00123	0,00077	5,5E-07	0	0,00108	0,0004	0,00204	0,00019	0,00041	0,00177	0,0025	0,00219
ы	0	0,00096	0,00147	0	0,0002	0	0	0	0	0,00036	0	0	0	0	0,00013	0,00142
й	0,00012	0	0	0	0	0,0021	0	3,3E-06	0	0	0,00111	0,00066	0	0	0	5,5E-07
к	0,00327	5E-05	0,00039	3,9E-05	0,00011	0,00169	0	0,00022	2,2E-05	8,6E-05	0,00139	0,0002	0,00024	0,00013	8,8E-05	0,00024
л	0,00216	0,00048	0,00141	0,00028	0,00054	0,00347	5E-06	9,7E-05	1,1E-06	0,00023	0,00095	0,0033	6,3E-05	0,00037	0,00012	0,00011
м	0,00137	2,4E-05	0,00021	8,8E-06	0,00023	0,00313	1,7E-06	1,7E-05	0	0,00071	0,00287	0,00112	0,0002	0,00011	1,1E-05	0,00018
н	0,00391	0,00031	0,00159	0,0001	0,0008	0,00825	5,5E-07	7E-05	0,00072	0,00148	0,00234	0,00023	0,00035	0,00023	0,00055	0,00055
о	0,00051	0,00124	0,00445	0,0039	0,00132	0,00133	0	5,5E-07	3,9E-06	0,00046	0,00122	0,0002	0,00041	0,00446	0,00288	0,00272
п	0,00128	1,5E-05	0,00053	8,8E-06	8,7E-05	0,00162	0	7,2E-06	5,5E-07	5,6E-05	0,00042	0,00063	0,00025	5E-05	0,0006	0,00018
р	0,00095	0,00068	0,00045	0,00054	0,00071	0,00349	5,5E-07	3,3E-06	5,5E-07	0,00031	0,00117	0,00028	0,00063	0,00025	1,7E-05	0,00017
с	0,00153	0,00016	0,00149	2,3E-05	0,00066	0,00482	0	2,8E-06	5,5E-07	5,5E-05	0,00287	0,00079	0,00176	0,00056	0,00014	0,00065
т	0,00227	2,8E-06	0,00054	3,3E-06	5,1E-05	0,00441	0	0,00186	0	2,4E-05	0,0032	0,00107	0,00027	0,00128	2,6E-05	0,0002
у	0,00022	0,0004	0,0007	0,00016	0,00066	0,00027	0	0	5,7E-05	0,00051	0,00024	7,1E-05	6,9E-05	0,00044	0,0006	0,001
ф	9,8E-05	0	7E-05	5,5E-07	3,3E-06	0,00023	0	7E-05	0</							

	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ь	ю	я	
а	0,00491	0,0001	0,00058	0,00486	0,00081	0,0033	0,0002	0,00022	0,00039	8,9E-05	0,0011	0,0002	0,00022	0	4,2E-05	2,1E-05	7,4E-05	
б	2,8E-05	0,00351	6,6E-06	2,1E-05	4,4E-05	0,00015	0,00023	0	5E-05	0	5,5E-07	0	1,1E-06	0	7,1E-05	4,9E-05	8,6E-05	
в	7,9E-05	0,00623	1E-05	0,00045	0,00129	0,00365	0,00048	5,5E-07	0,00033	2E-05	6,6E-06	3,3E-06	2,8E-06	0	0,0004	0,00013	0,00129	
г	5,2E-05	0,00331	1,1E-06	0,00039	1,7E-05	3,9E-05	0,00052	5,5E-07	3,2E-05	2,2E-06	1,1E-06	5,5E-07	0	0	3,3E-05	8,3E-06	6,6E-05	
д	0,00018	0,00294	6,1E-06	3,1E-05	0,0002	0,00034	0,00062	0	0,00019	3,9E-06	1,7E-06	0	0	0	0,00014	0,00028	0,00045	
е	0,00372	0,00153	0,00165	0,00462	0,00102	0,00417	0,00066	0,00019	7,9E-05	0,00093	0,00329	0,00111	0,0017	0,00029	0,00019	2,4E-05	0,00072	
ё	5,5E-07	5,5E-07	0	0	1,1E-06	5,5E-07	0	0	0	0	5,5E-07	1,1E-06	5,5E-07	0	0	0	0	
э	9,9E-06	0,00034	6,6E-06	1E-05	7,1E-05	8,3E-05	5E-05	2E-05	0,00012	3,3E-06	0	0	0	0	9,2E-05	2,3E-05	0,00019	
ж	3,3E-06	0,00125	5,5E-07	0,00023	1,4E-05	3,4E-05	0,00057	0	8,2E-05	1,1E-06	0	0	0	0	4,1E-05	1,9E-05	0,00011	
з	1,4E-05	0,00166	1,7E-06	1,3E-05	2,2E-05	7,1E-05	8,3E-05	0	0,0001	0	1,1E-06	0	0	0	0,00015	3,1E-05	0,00059	
и	0,00748	0,00124	0,00037	0,0028	0,00132	0,0043	0,0002	0,00032	0,0006	0,00138	0,00095	0,00052	0,00086	0	0,00026	9,9E-05	0,00061	
ы	0,00327	0	0,00078	0,00091	0,00017	0,00045	0	5,5E-07	0	2,8E-05	0	0	0	0	0	0	0	
й	5,5E-07	0,00262	0	0	0	0	2,8E-06	0	0	0	0	0	0	0	0	1,7E-06	5,5E-07	
к	0,00073	0,00111	1,2E-05	0,00019	0,00189	0,00039	0,00051	5,5E-07	0,00015	5,5E-06	0,00014	0,00022	0	0	0,00055	0,00013	0,0005	
л	1,8E-05	0,00277	0,00025	1,9E-05	0,00221	0,00018	0,00032	0,0001	7,2E-05	5,5E-07	1,3E-05	7,8E-05	0	0	4,7E-05	3,1E-05	0,00018	
м	1,5E-05	0,00299	6,1E-06	0,00055	0,00041	0,00018	0,00028	7,7E-06	0,00012	5E-06	1,7E-06	5,5E-07	0	0	0,00017	4,8E-05	0,00054	
н	0,00278	0,00255	6,2E-05	0,00033	0,00064	0,00166	0,00051	1,2E-05	0,00024	7,2E-06	0,00068	0,00021	4,6E-05	0	0,00293	7,8E-05	0,00079	
о	0,00847	0,00141	0,00407	0,00393	0,00285	0,00706	0,00018	0,0005	0,00159	1,5E-05	9,9E-06	5,5E-05	0	0	0,00027	0,00015	0,00053	
п	6,7E-05	0,00257	0,00012	8E-05	0,00155	0,00043	0,0006	1,1E-06	0,00043	9,4E-06	5,5E-06	1,7E-06	0	0	0,00037	0,00015	0,00071	
р	4E-05	0,00391	0,00494	2,5E-05	0,00058	0,00177	0,00035	2,6E-05	0,00027	3,9E-06	2E-05	5,5E-07	0	0	0,0001	5,7E-05	0,0003	
с	0,00033	0,0057	0,00057	0,00012	0,00125	0,00224	0,00097	3E-05	0,00055	9,4E-06	3,9E-06	2,2E-06	0	0	0,00065	0,00021	0,00083	
т	0,00083	0,00479	7,4E-05	0,00027	0,00867	0,00018	0,00064	1,7E-06	0,00011	3,3E-06	0,0011	1,8E-05	5,5E-07	0	0,00037	0,00084	0,00149	
у	0,00068	0,00024	0,00024	0,00144	0,00061	0,00102	3,9E-05	0,00023	0,00015	2E-05	0,0003	3,4E-05	0,00023	0	8,6E-05	2,5E-05	0,00019	
ф	1,7E-05	0,00016	1,7E-06	2,8E-05	9,2E-05	4,5E-05	2,6E-05	0,00014	0,00011	0	1,1E-06	0	0	0	2,6E-05	4,4E-05	9,4E-05	
х	6,1E-06	0,00016	1,7E-06	5,3E-05	0,00025	2,6E-05	0,00012	0	8,8E-06	0	5,5E-07	0	0	0	1E-05	8,3E-06	0,00033	
ц	0,00024	0,0007	3E-05	2,2E-05	1E-05	1,1E-05	1E-05	0	2,1E-05	0	0	2,8E-06	5,5E-07	0	1,8E-05	4,1E-05	5,1E-05	
ч	3,4E-05	0,00104	7,2E-06	2,4E-05	0,0001	0,00016	0,00091	0	6,5E-05	2,2E-06	1,1E-06	0	0	0	0,00013	0,00023	0,0002	
ш	1,4E-05	0,00065	7,2E-06	0,00016	0,00011	5,5E-06	9E-05	0	5,5E-06	0	1,2E-05	2,2E-06	0	0	0,00018	2,8E-06	8,3E-06	
щ	9,4E-06	0,00026	0	2,2E-06	0	5,5E-06	0,00064	0	0	0	0	0	0	0	2,2E-06	0,00099	0,00017	
ъ	0	0	0	0	5,5E-06	0	0	0	0	0	0	0	0	0	0	0	0	
ь	0,00026	0	4,4E-06	0,00013	0,00048	0,00246	0	0	0	0	0	5E-05	0,00018	4,5E-05	0	0	0	0
ю	5,3E-05	5,4E-05	0	8,8E-06	2,9E-05	1,1E-06	0,00107	0	5,5E-07	0	0	0	0	0	0,00023	2E-05	0,00032	
я	0,00049	0,00038	3,3E-05	0,00032	0,00225	0,00018	4,5E-05	0	4,4E-05	5,5E-07	0	0	0	0	8,8E-05	0,0001	4,4E-06	9,8E-05

Ентропія:
H2: 4.5819706093794315
R: 0.09167077334509166

Частота біграм без перетину:

а	б	в	г	д	е	ё	з	ж	з	и	ы	й	к	л	м	
а	8,9E-05	0,00012	0,00301	0,00062	0,00184	0,00036	0	0	0,00055	0,00173	0,0005	3,2E-05	9E-05	0,00285	0,00114	0,00107
б	0,00047	1,1E-06	6,1E-05	4,4E-06	1,8E-05	0,00093	4,4E-06	1,1E-06	2,2E-06	5,5E-05	0,00041	0,00013	6,4E-05	0,00011	5,5E-06	0,0001
в	0,00192	8,8E-06	0,00018	1,9E-05	0,00052	0,00147	1,1E-06	3,1E-05	0	0,00111	0,00271	0,00079	0,00038	0,00023	3,8E-05	0,00043
г	0,00026	1,1E-06	3,3E-05	4,4E-06	4,1E-05	0,00139	0	4,4E-06	0	0,00011	0,00057	5,6E-05	7E-05	7,7E-06	1,8E-05	5E-05
д	0,00099	9,9E-06	0,00025	0,0003	3,6E-05	0,00305	0	1,1E-06	0,00055	0,00072	0,00094	0,00018	0,00034	0,0001	2,8E-05	0,00022
е	0,00176	0,00094	0,00303	0,00022	0,00362	0,00107	0	0	0,0017	0,00012	0,00306	0,00121	8,7E-05	0,00019	0,00301	0,00275
ё	0	0	1,1E-06	0	0	0	0	0	0	0	0	0	0	7,7E-06	0	0
э	0,00019	3,2E-05	0,00018	1,1E-06	9,9E-06	0,00026	0	0	0	3,3E-05	0,00027	5,1E-05	0,00012	5,8E-05	1,5E-05	0,00013
ж	0,00086	1,1E-06	2E-05	1,1E-06	7,7E-06	0,00055	0	0	0	1,1E-05	0,00033	2,2E-05	9,5E-05	0,0002	0,00014	5,4E-05
з	0,00235	6,6E-06	0,0003	1,1E-06	1,8E-05	0,00087	0	0	0	1,2E-05	0,00229	0,00013	0,0001	4E-05	8,8E-06	9,5E-05
и	0,00062	0,0002	0,00314	0,00086	0,00126	0,00076	1,1E-06	0	0,0011	0,00041	0,00202	0,00018	0,00039	0,00176	0,0025	0,00215
ы	0	0,00097	0,0015	0	0,00021	0	0	0	0	0,00038	0	0	0	0	0,00013	0,00146
й	0,00011	0	0	0	0	0,00204	0	5,5E-06	0	0	0,00109	0,00068	0	0	0	1,1E-06
к	0,00325	5,3E-05	0,00039	3,4E-05	0,00012	0,00169	0	0,00022	2E-05	8,2E-05	0,00144	0,00002	0,00025	0,00014	0,0001	0,00023
л	0,00213	0,00051	0,00146	0,00028	0,00052	0,0035	1,1E-06	9,4E-05	1,1E-06	0,00022	0,00096	0,00032	7,1E-05	0,00036	0,00012	0,00011
м	0,00135	2E-05	0,00023	5,5E-06	0,00022	0,00313	0	1,8E-05	0	0,00068	0,00293	0,00113	0,00021	0,00011	7,7E-06	0,00019
н	0,00388	0,00032	0,00163	0,0001	0,0008	0,00826	0	7E-05	0,00072	0,00148	0,00231	0,00024	0,00035	0,00032	0,00024	0,00062
о	0,00052	0,00124	0,00452	0,00392	0,0013	0,00127	0	0	5,5E-06	0,00044	0,00123	0,0002	0,00042	0,00448	0,00286	0,00269
п	0,00134	1,2E-05	0,00052	1,1E-05	8,6E-05	0,00162	0	8,8E-06	0	6,4E-05	0,00116	0,00041	0,00065	0,00024	5,3E-05	0,00058
р	0,00097	0,00068	0,00043	0,00052	0,00071	0,00351	0	5,5E-06	1,1E-06	0,00031	0,00112	0,00029	0,00025	0,00058	1,9E-05	0,00017
с	0,00149	0,00016	0,00145	2,3E-05	0,00064	0,00482	0	4,4E-06	0	5,5E-05	0,00298	0,00077	0,00183	0,00057	0,00014	0,00061
т	0,00228	4,4E-06	0,00057	1,1E-06	5,6E-05	0,00445	0	0,00187	0	2,1E-05	0,00321	0,00108	0,00024	0,00127	2,8E-	

	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ь	ю	я	
а	0,00496	0,00011	0,00059	0,00497	0,00081	0,00333	0,0002	0,00021	0,00039	8,2E-05	0,0011	0,00019	0,00023	0	3,6E-05	1,9E-05	6,1E-05	
б	3,1E-05	0,00345	5,5E-06	2,4E-05	4,6E-05	0,00015	0,00022	0	4,2E-05	0	0	0	2,2E-06	0	6,4E-05	5,2E-05	8,5E-05	
в	8,1E-05	0,00621	1,3E-05	0,00045	0,0013	0,00372	0,00051	0	0,00036	2,6E-05	7,7E-06	2,2E-06	3,3E-06	0	0,00039	0,00013	0,0012	
г	4,6E-05	0,0034	0	0,00042	2E-05	4,4E-05	0,00054	0	3,2E-05	2,2E-06	1,1E-06	0	0	0	3,3E-05	8,8E-06	5,8E-05	
д	0,00017	0,00291	3,3E-06	2,8E-05	0,0002	0,00036	0,00065	0	0,00019	3,3E-06	2,2E-06	0	0	0	0,00014	0,00028	0,00043	
е	0,00367	0,00153	0,00167	0,00462	0,00104	0,00411	0,00065	0,00021	9E-05	0,00092	0,00332	0,00114	0,0017	0,00028	0,0002	2,5E-05	0,00072	
ё	1,1E-06	1,1E-06	0	0	1,1E-06	0	0	0	0	1,1E-06	2,2E-06	0	0	0	0	0	0	
э	1,1E-05	0,00036	4,4E-06	8,8E-06	6,3E-05	8,1E-05	5,5E-05	2,1E-05	0,00013	3,3E-06	0	0	0	0	8,5E-05	2,4E-05	0,00018	
ж	3,3E-06	0,00122	1,1E-06	0,00025	1,7E-05	3,2E-05	0,00056	0	7,2E-05	1,1E-06	0	0	0	0	3,4E-05	2E-05	0,00011	
з	1,1E-05	0,00167	2,2E-06	1,7E-05	2,9E-05	8,5E-05	8,9E-05	0	0,0001	0	1,1E-06	0	0	0	0,00015	3,1E-05	0,00058	
и	0,00742	0,00125	0,00038	0,0028	0,00135	0,00436	0,0002	0,00034	0,00062	0,00142	0,00094	0,00053	0,00086	0	0,00027	0,0001	0,00062	
ы	0,00322	0	0,00077	0,00091	0,00016	0,00045	0	0	0	2,8E-05	0	0	0	0	0	0	0	
й	0	0,00261	0	0	0	0	2,2E-06	0	0	0	0	0	0	0	0	2,2E-06	1,1E-06	
к	0,00077	0,00107	1,3E-05	0,00022	0,00191	0,0004	0,00051	0	0,00014	5,5E-06	0,00013	0,00022	0	0	0,00056	0,00013	0,0005	
л	1,9E-05	0,00276	0,00025	1,9E-05	0,00221	0,00019	0,00031	8,4E-05	7,9E-05	1,1E-06	1,3E-05	7,8E-05	0	0	4,7E-05	2,8E-05	0,00018	
м	1,2E-05	0,00302	5,5E-06	0,00054	0,00039	0,00017	0,00027	6,6E-06	0,00013	5,5E-06	1,1E-06	0	0	0	0,00017	4,7E-05	0,00055	
н	0,00281	0,00256	6,1E-05	0,00031	0,00068	0,00166	0,00052	1,1E-05	0,00024	1,1E-05	0,00067	0,00021	4,4E-05	0	0,00296	8,5E-05	0,0008	
о	0,00843	0,00146	0,00409	0,00401	0,00285	0,00699	0,0002	0,0005	0,00162	1,4E-05	8,8E-06	6,1E-05	0	0	0,00026	0,00017	0,00052	
п	7,5E-05	0,00257	0,00013	6,3E-05	0,00151	0,00042	0,00062	0	0,00043	7,7E-06	4,4E-06	1,1E-06	0	0	0,00037	0,00015	0,00075	
р	3,8E-05	0,00388	0,00486	2,2E-05	0,00058	0,00181	0,00033	2,9E-05	0,00027	5,5E-06	1,9E-05	1,1E-06	0	0	9,4E-05	6E-05	0,0003	
с	0,00033	0,0057	0,00056	0,00012	0,00126	0,00221	0,00101	3,3E-05	0,00056	9,9E-06	1,1E-06	3,3E-06	0	0	0,00066	0,0002	0,00086	
т	0,00082	0,00477	7,8E-05	0,00023	0,00853	0,00018	0,00064	2,2E-06	0,00011	2,2E-06	0,00114	1,9E-05	1,1E-06	0	0,00035	0,00085	0,00148	
у	0,00065	0,00023	0,00024	0,00142	0,0006	0,00099	4,4E-05	0,00023	0,00015	2,5E-05	0,00028	3E-05	0,00024	0	7,8E-05	2,2E-05	0,00018	
ф	1,7E-05	0,00016	2,2E-06	2,2E-05	8,7E-05	4,2E-05	2,6E-05	0,00014	9,7E-05	0	1,1E-06	0	0	0	2,2E-05	4,4E-05	0,0001	
х	6,6E-06	0,00015	1,1E-06	4,7E-05	0,00024	2,6E-05	0,00012	0	8,8E-06	0	0	0	0	0	1,4E-05	7,7E-06	0,00033	
ц	0,00022	0,00071	2,9E-05	2E-05	8,8E-06	1,4E-05	9,9E-06	0	2E-05	0	0	1,1E-06	0	0	0	1,7E-05	4,3E-05	5,6E-05
ч	3,2E-05	0,00103	8,8E-06	2,6E-05	0,0001	0,00015	0,00096	0	6,3E-05	2,2E-06	1,1E-06	0	0	0	0,00014	0,00022	0,00019	
ш	1,2E-05	0,00063	4,4E-06	0,00014	0,00012	3,3E-06	8,6E-05	0	5,5E-06	0	1,2E-05	2,2E-06	0	0	0	0,00017	1,1E-06	1,1E-05
щ	7,7E-06	0,00025	0	2,2E-06	0	6,6E-06	0,00062	0	0	0	0	0	0	0	0	1,1E-06	0,00098	0,00018
ъ	0	0	0	0	6,6E-06	0	0	0	0	0	0	0	0	0	0	0	0	
ь	0,00025	0	5,5E-06	0,00014	0,00048	0,00251	0	0	0	0	4,6E-05	0,00019	4,5E-05	0	0	0	0	
ю	5,7E-05	5,3E-05	0	9,9E-06	2,1E-05	2,2E-06	0,00111	0	0	0	0	0	0	0	0,00021	1,7E-05	0,00032	
я	0,00048	0,00038	2,9E-05	0,00033	0,00226	0,00017	5,6E-05	0	4E-05	1,1E-06	0	0	0	9,9E-05	9,6E-05	2,2E-06	0,0001	

Ентропія:

H2: 4.581302570504652

Надлишковість:

R: 0.09180320528022046

Використовуючи програму CoolPinkProgram, оцінимо значення H⁽¹⁰⁾, H⁽²⁰⁾, H⁽³⁰⁾:

Лабораторная работа №1

Произвольная часть текста:
орить_здесь_не_имеет_к_ним_отношения_итак_возвратимся_к_обычным_человечески

Использованные буквы:

Порядок n-грамм:
5 символов
10 символов
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: с

Символ по счету: 1

Номер эксперимента: 51

Неравенство для энтропии:
2,20463740567953 < H < 2,86397470118086

Поле ввода символов:
с

Продолжить Другой

Двоичная таблица угаданных символов:

00100000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
00000000100000000000000000000000
10000000000000000000000000000000

Строка состояния:
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

Вероятности:

```

q[1] = 0,5098039
q[2] = 0,0588235
q[3] = 0,0588235
q[4] = 0,0196078
q[5] = 0
q[6] = 0,0588235
q[7] = 0,0196078
q[8] = 0,0196078
q[9] = 0,0196078
q[10] = 0
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0,039215
q[17] = 0
q[18] = 0,019607
q[19] = 0,019607
q[20] = 0,019607
q[21] = 0,019607
q[22] = 0,019607
q[23] = 0
q[24] = 0
q[25] = 0,019607
q[26] = 0
q[27] = 0,058823
q[28] = 0,019607
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

```

2.20463740567953 < H⁽¹⁰⁾ < 2.86397470118086

$$1.14417636367381 < H^{(20)} < 1.78059178105309$$

$2.96693780180905 < H^{(30)} < 3.30036092507113$

Обчислимо надлишковість для $H^{(10)}$, $H^{(20)}$, $H^{(30)}$:

-для $H^{(10)}$: $0.432246047113959 < R < 0.562952982357390$;

-для $H^{(20)}$: $0.647015728961411 < R < 0.7731786342223144$;

-для $H^{(30)}$: $0.345736901800435 < R < 0.411834656133810$.

Висновки:

У результаті виконання комп'ютерного практикуму ми навчилися порівнювати різні моделі відкритого тексту, засвоїли поняття ентропії на символ джерела та його надлишковості за допомогою практичних навичок роботи з біграмами і літерами.