

Міністерство освіти і науки України  
Національний технічний університет України  
“Київський політехнічний інститут імені Ігоря Сікорського”  
Фізико-технічний інститут

Криптологія  
Комп’ютерний практикум

Робота №1

Виконали:  
студенти групи ФБ-14  
Мартиненко Даніїл  
Цуканов Данило

Київ-2023

## Мета: оцінити надлишковість російської мови в різних моделях джерела

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 ймовірності замінити відповідними частотами. Також одержати значення  $H_1$  та  $H_2$  на тому ж тексті, в якому вилучено всі пробіли.

Код на python:

```
import math
no_target_symb='\n '

file_name = "res1.txt"

def counf_symb():
    with open(file_name, "r", encoding="cp1251") as f:
        count = 0
        while True:
            c = f.read(1).lower()
            if not c:
                break
            if c not in no_target_symb:
                count += 1
        return count

def count_c1():
    c1 = dict()
    with open(file_name, "r", encoding="cp1251") as f:
        while True:
            c = f.read(1).lower()
            if not c:
                break
            if c not in no_target_symb:
                if c not in c1.keys():
                    c1[c] = 0
                c1[c] += 1
    return c1

def count_c2():
    c2 = dict()
    with open(file_name, "r", encoding="cp1251") as f:
        a = f.read(1).lower()
        while True:
            b = f.read(1).lower()
            if not b:
                break
            if b not in no_target_symb:
                if a+b not in c2.keys():
                    c2[a+b] = 0
                c2[a+b] += 1
            a = b
    return c2

def count_c2_nocross():
    c2_nocross = dict()
    with open(file_name, "r", encoding="cp1251") as f:
        text = f.read().replace('\n', '').replace(' ', '')
        i = 0
        while True:
            if (i >= len(text) - 1):
                break
            b = text[i] + text[i + 1]

            if b not in c2_nocross.keys():
                c2_nocross[b] = 0
            c2_nocross[b] += 1
            i += 2
        return c2_nocross

def count_h(p):
    res = 0
    for i in p.keys():
        res += p[i] * math.log2(p[i])
    return -res

count_of_symbols = counf_symb()
c1 = count_c1()
c2 = count_c2()
c2_nocross = count_c2_nocross()
p1 = dict()
p2 = dict()
p2_nocross = dict()
for i in c1.keys():
    p1[i] = c1[i] / count_of_symbols

for i in c2.keys():
    p2[i] = c2[i] / count_of_symbols

for i in c2_nocross.keys():
    p2_nocross[i] = c2_nocross[i] / (count_of_symbols / 2)

h1 = count_h(p1) / 1
h2 = count_h(p2) / 2
h2_nocross = count_h(p2_nocross) / 2
print("*****h1*****")
print(h1)
print("\n*****h2*****")
print(h2)
print("\n*****h2_nocross*****")
print(h2_nocross)
print("*****p2_nocross*****")
for i in p2_nocross.keys():
    print(f"{i}:{p2_nocross[i]}", end=',')
```

Із пробілами:

H1 і H2

```
*****h1*****  
4.387662336862179  
  
*****h2*****  
3.9909063611403606  
  
*****h2_nocross*****  
3.989031146917921
```

3.989031146917921

Частоти букв:

' ': 0.15963314117636226  
а: 0.06611209835579815  
б: 0.014299190663966926  
в: 0.03499673135049214  
г: 0.01472943031749845  
д: 0.02932493389459798  
е: 0.0691857170479533  
ж: 0.008854231624623435  
з: 0.014888468321722165  
и: 0.052823217539715556  
й: 0.009239271003270324  
к: 0.030726142416021656  
л: 0.045843960280676964  
м: 0.025580844458320744  
н: 0.05515019465414675  
о: 0.09526209044575004  
п: 0.02290482077672487  
р: 0.038152380171141635  
с: 0.04245310262220187  
т: 0.051782774175241464  
у: 0.025307131682630456  
ф: 0.0007784491785687082  
х: 0.006693825893563397  
ц: 0.002591482426719054  
ч: 0.013082968452719257  
ш: 0.007274733130043702  
щ: 0.00319080459000421  
ь: 0.01695847350301293  
ы: 0.01612477954402967  
э: 0.004500775519531122  
ю: 0.004943570804975044  
я: 0.01661026397797575

Частоти біграм з пересіченнями 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	ь	ы	э	ю	я		
	0	0,0168	0,00025	0,00552	0,0007	0,00192	0,01608	0,0003	0,00164	0,01782	0,00709	0,0058	0,00802	0,00697	0,00377	0,02116	7,7E-05	0,00101	0,00249	0,00531	0,00622	1,8E-05	0,00282	0,00032	9E-05	6,2E-05	1,3E-05	0,01041	0,00453	1,3E-05	0,00249	0,00983	
а	0,00156	3E-05	0,00101	0,00483	0,00107	0,00426	6,9E-05	0,00121	0,0057	9,2E-05	1,7E-05	0,00773	0,00688	0,00304	0,0095	5,9E-05	0,00168	0,0067	0,00122	0,00573	8,9E-05	0,00012	0,0004	0,00066	0,00167	0,00084	0,0004	1,7E-06	8,4E-07	0	1,1E-05	1,7E-06	
б	0,00679	0,00067	4,2E-06	1,3E-05	8,4E-07	2,8E-05	0,00124	8,4E-06	0,00013	0,00051	6,7E-06	2,5E-06	3,3E-05	3,7E-05	8,4E-07	0,00326	0	3,4E-05	6,5E-05	0,00014	0,0008	4,2E-06	0	0	0	5E-06	0	4,7E-05	0,00028	1,7E-06	0,00016	2,2E-05	
в	0,01384	0,00272	3,4E-05	8,7E-05	1,3E-05	0,00123	0,00172	0	0,00083	0,00204	1,6E-05	0,00017	3,4E-05	1,3E-05	5E-05	0,00578	1,7E-06	0,00047	0,00121	0,00201	0,00088	0	0,00026	6,9E-05	1,1E-05	2,9E-05	0	0,00029	0,00091	5E-06	1,1E-05	0,00026	
г	0,00361	0,00071	5E-06	1,5E-05	7,5E-06	2,2E-05	0,00319	1,5E-05	0,00042	0,00055	3,3E-06	5E-06	0,00015	6,9E-05	7,7E-05	0,00436	8,4E-07	0,00013	4,4E-05	1,5E-05	0,00101	0	6,7E-06	0	1,7E-06	0	0	2,4E-05	0,00013	7,5E-06	2,2E-05	0,00013	
д	0,00724	0,00198	8,4E-06	0,00026	0,00129	0,00164	0,00245	0,0006	0,00008	0,00152	0,00016	2,9E-05	8,2E-05	8,4E-06	0,0014	0,00463	8,4E-07	0,00037	0,00003	0,00001	0,00175	0	2,5E-06	1,7E-06	0	0	0	6,4E-05	0,0001	0,00163	0,00026	0,00064	
е	0,00363	0,00112	0,002	0,00512	0,00037	0,00459	0,00131	0,00363	0,00004	0,00179	6,7E-06	0,0007	0,00408	0,00348	0,00909	0,00192	0,00204	0,00727	0,00287	0,00433	0,00014	0,00011	6,9E-05	0,00071	0,00347	0,00193	0,00145	0,00039	0,00107	0	8,4E-07	9,9E-05	
ж	0,00166	0,00114	1,9E-05	3,3E-06	0	0,00074	0,00072	1,6E-05	0,0001	0,00033	1,7E-06	8,4E-06	0,00025	8,4E-07	1,5E-05	0,00207	0	0,0003	4,9E-05	3,3E-06	0,00117	0	0	0	0	0	0	0	4,6E-05	0	3,9E-05	0,00016	
з	0,00466	0,00355	9,2E-06	0,00055	3,3E-06	1,6E-05	0,00137	1,7E-06	8E-05	0,0021	1,3E-05	1,5E-05	6,3E-05	4,2E-06	9,9E-05	0,00135	8,4E-07	3,5E-05	1E-05	1,1E-05	0,00024	0	8,4E-07	0	0	0	0	0,00013	7,7E-05	1,7E-06	0,00028	0,00023	
и	0,00846	0,00011	0,00073	0,00282	0,00059	0,00038	0,00012	0,00137	0,00044	0,0004	3,3E-06	0,00286	0,00653	0,00287	0,00069	0,00088	0,00389	0,00146	0,00318	3,9E-05	0,00013	0,00022	0,00026	0,00184	0,00139	0,001	4E-05	8,4E-06	0	1,7E-06	1,2E-05		
й	4,2E-06	0,00057	0	0	0	0	0,00204	0	0	0,00104	8,4E-07	0	0	0	0,00379	0	0	0	0	0	0,00014	0	0	0	0	0	0	0	0,00148	6,9E-05	5,9E-05	3,8E-05	
к	0,00899	0,00415	0,00015	0,00023	0,00013	0,00025	0,00163	0,0001	0,00017	0,00197	0,00034	3,9E-05	0,00076	0,00016	0,00036	0,00331	0,00013	0,00038	0,00304	0,00074	0,00109	0	2,5E-06	1,8E-05	0,0005	0,00058	0	0,00098	0,00017	6,9E-05	9,1E-05	0,00019	
л	0,00242	0,00905	0,00069	0,00055	0,0017	0,00055	0,00713	7,5E-06	0,00027	0,00489	3,1E-05	0,00075	0,00022	0,00026	4,2E-06	0,00727	0,00081	0,00016	0,00278	0,00021	0,00022	5,4E-05	0,00017	0	1,1E-05	0,0004	0	0	0,00224	7E-05	1,3E-05	0,00093	
м	0,00622	0,00268	4,2E-05	0,00012	8,4E-06	5E-05	0,00357	3,3E-06	0,00042	0,00237	9,9E-05	5E-06	8E-05	4,4E-05	3,3E-06	0,0051	0	0,00024	0,00097	5,2E-05	0,0014	5E-06	5,5E-05	0	3,3E-06	3E-05	8,4E-07	0,00002	0,00135	2,2E-05	6,8E-05	0,00035	
н	0,01508	0,00416	0,00037	0,00129	0,00043	0,00188	0,00617	0,00098	0,0015	0,00329	0,00037	0,00042	0,00043	0,00126	0,00286	0,00666	0,00021	0,00127	0,00102	0,00102	0,00004	0	0,00023	8,4E-07	0,00091	0,00044	7E-05	0,00126	0,00019	0,00013	9,7E-05	0,00075	
о	0,01125	6,4E-05	0,00198	0,00569	0,00703	0,00325	0,00018	0,00024	0,00052	0,00018	5,2E-05	0,00798	0,00754	0,0037	0,00999	0,0002	0,00943	0,00792	0,00231	0,01283	0,0001	0,00017	0,00196	0,00024	0,00012	0,00026	2,5E-06	3E-05	4,2E-06	8,4E-07	3,3E-06	7,5E-06	
п	0,01592	0,00092	0	0,00031	1,7E-06	8,5E-05	0,00097	0	1,8E-05	0,00027	1,5E-06	5E-06	3,7E-05	9E-05	7,5E-06	0,0014	2,2E-05	9,6E-05	0,0017	0,00001	0,00065	0	5E-06	8,4E-07	0	1,5E-05	0	0	0,00019	5E-06	5E-06	4,4E-05	
р	0,00522	0,00223	0,00118	0,0007	0,00083	0,00108	0,00648	1,8E-05	0,0003	0,001	6,9E-05	0,00191	4,2E-06	8,5E-05	0,00026	0,00504	0,00565	2,7E-05	0,00017	0,00046	0,00064	3,2E-05	0,00018	8,4E-07	1,4E-05	2,3E-05	5E-06	1,3E-05	0,00035	2,8E-05	2,3E-05	0,00012	
с	0,01576	0,00337	0,00017	0,00012	2,3E-05	0,0002	0,00377	3,3E-06	1,7E-06	0,00251	0,00047	0,0002	0,00013	0,00031	0,00581	3,6E-05	0,00021	0,00061	0,00118	0,00139	2,5E-06	4,9E-05	0	1,7E-06	0	0	0,0009	0,00065	1,2E-05	0,00013	0,00065		
т	0,00826	0,00403	3,3E-06	0,00023	2,4E-05	0,00021	0,00539	4,2E-06	2,1E-05	0,00307	0,00023	0,00045	0,00015	4,9E-05	0,00054	0,00665	0,00012	0,00112	0,01068	0,00051	0,00163	7,5E-06	2,5E-05	1,7E-06	0,00334	0,00014	0	8,9E-05	0,00076	0,00238	0,00032	0,00135	
у	0,00452	0,00017	0,0011	0,00068	0,00047	0,00167	9,5E-05	0,00023	0,00041	7,5E-06	0,00197	0,00012	0,00185	0,00351	7,4E-05	0,00098	0,00268	0,00073	0,00138	1,6E-05	1E-04	0,00016	0,00012	0,00077	0,00023	0,00023	8,4E-07	6,7E-06	0	8,4E-07	4,2E-06	0	
ф	0,00043	4,3E-05	0	0	0	0	4,4E-05	0	0	3,1E-05	1E-05	0	5,9E-06	2,5E-06	1,7E-05	8,2E-05	0	2,5E-05	3,2E-05	1,4E-05	1,3E-05	9,2E-06	8,4E-07	0	0	0	0	9,2E-06	8,4E-07	5E-06	0	8,4E-07	
х	0,00135	0,00089	4,7E-05	3,1E-05	8,4E-07	5,3E-05	0,00041	0	0	0,00137	2,5E-06	1,7E-06	5,9E-06	5,9E-06	0,00061	0	0,00019	0,0002	1,3E-05	0,00048	0	7,5E-06	0	0	0	0	0	4,2E-06	0,00083	1,3E-05	3,9E-05	0,00012	
ц	0,00027	0,00017	6,7E-06	3,4E-05	0	0,00022	0,00028	0	3,3E-06	0,00075	1,8E-05	2,7E-05	0	5E-06	0,0003	8,7E-05	5,9E-06	4,9E-05	4,1E-05	8E-05	9,2E-06	0	1,7E-06	1E-05	0	9,2E-06	0	0,00017	1,1E-05	0	3,3E-06	2,1E-05	
ч	0,00558	0,00088	5,9E-06	4,4E-05	2,9E-05	1,8E-05	0,00082	7,1E-05	4,2E-06	0,00085	0,00016	6,7E-06	0,00012	5,3E-05	0,0002	0,00192	2,5E-05	0,00014	0,00029	0,00013	0,00048	2,5E-06	1,7E-06	0	3,3E-06	0	0	0,00046	0,00017	2,3E-05	0,00016	0,00013	
ш	0,00092	0,00068	8,4E-06	0,00088	5,9E-06	5,9E-05	0,00096	0	2,5E-06	0,00073	6,5E-05	2,3E-05	2,4E-05	3,3E-06	2,5E-06	0,00089	4,2E-06	0,00017	8E-05	2,9E-05	0,00049	8,4E-07	1,5E-05	0	9,4E-05	6,7E-06	0	0,00055	0,00052	6,7E-06	4,4E-05	2,4E-05	
щ	0,00013	0,00034	9,3E-05	9,2E-06	0	0	0,0007	0	0	0,00019	0	3,3E-06	8,4E-06	1,7E-06	0,00022	0,00025	8,4E-07	2,3E-05	3,3E-06	1,1E-05	0,00031	0	0	0	0	0	0	5E-06	1,6E-05	0	0,00058	0,0003	
ь	0	0	0,00018	0,00028	0	0,00049	0	4E-05	0,00017	0	0	8,4E-07	0,00418	5,4E-05	0,00099	0	3,7E-05	0,00688	0,00308	0,00566	0	8,4E-07	1,7E-06	0	0,00023	0,00087	2E-05	0	0	0	0	0	
ы	0	0	0,00374	0,00223	2,5E-06	0,00064	0	0	0,00037	0	0	0	0,00087	0,00076	0,00327	0	0,0004	0,00146	0,00034	0,00184	0	1E-05	0	0,00018	0	0	0	0	0	0	0	0	0
э	0,00412	2,5E-05	1,7E-06	5E-05	8,4E-07	3,4E-05	8,4E-06	0	0	1E-05	2,5E-06	1,4E-05	1,3E-05	3,3E-05	3,3E-06	5,1E-05	1,7E-06	1,2E-05	1,8E-05	5E-06	1,6E-05	8,4E-07	4,8E-05	0	4,2E-06	7,5E-06	0	1,1E-05	1,7E-06	2,5E-06	2,5E-06	4,2E-06	
ю	0,00011	0,00097	3,3E-06	2,5E-06	0	8,5E-05	9,8E-05	0	0	0,00019	0	8,4E-07	0,00075	0	0,00021	0,00016	1E-05	0,00016	0,00037	5,7E-05	0,00112	0	0	0	0	8,4E-07	0	0,00052	0	0	3,6E-05	1E-04	
я	0,00153	0,00187	0,00045	0,00028	0	0,00025	0,00016	0	0,00016	0,00092	1,7E-06	8,4E-07	0,00163	0,00053	0,00136	0,00066	0,0004	0,00094	0,00428	0,00063	7,4E-05	8,4E-07	0	0	0	0	0	0,00036	1,7E-05	1,7E-06	0	0,0001	

Частоти біграм без пересічень 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	ь	ы	э	ю	я		
	0	0,016991	0,000248	0,005535	0,000732	0,001955	0,015924	0,000296	0,016399	0,017976	0,007016	0,005715	0,008999	0,006952	0,003701	0,021182	7,03E-05	0,000994	0,002498	0,00531	0,006156	1,84E-05	0,002797	0,00031	8,87E-05	6,70E-05	5,02E-06	0,010322	0,004423	1,00E-05	0,002525	0,0099	
а	0,001646	3,18E-05	0,000983	0,004831	0,001031	0,00417	7,03E-05	0,00119	0,005851	9,37E-05	0	0,007316	0,006911	0,003038	0,00941	5,02E-06	0,001639	0,006787	0,001158	0,000548	0,000102	0,000124	0,000429	0,000681	0,001696	0,000864	0,000412	0	1,67E-06	0	1,17E-05	1,67E-06	
б	0,006884	0,000773	5,02E-06	1,34E-05	1,67E-06	2,89E-05	0,001251	1,17E-05	0,000141	0,000509	3,35E-06	3,35E-06	3,01E-05	4,35E-05	1,67E-06	0,000246	0	3,01E-05	5,86E-05	0,000157	0,000782	3,35E-06	0	0	0	4,35E-05	0,000246	0	0,000142	2,68E-05	0		
в	0	0,002756	5,02E-05	8,87E-05	1,51E-05	0,001229	0,001708	0	0,000794	0,001952	6,70E-06	0,000167	3,85E-05	1,34E-05	4,52E-05	0,005769	1,67E-06	0,000445	0,001202	0,002061	0,000892	0	0,000273	6,33E-05	1,17E-05	2,34E-05	0	0	0,000275	0,000954	6,70E-06	6,70E-06	0,00026
г	0,003723	0,000715	6,70E-06	2,01E-05	8,87E-06	2,34E-05	0,002393	0,00151	5,15E-05	0,000407	0,000578	1,67E-06	5,02E-06	0,000146	6,53E-05	7,70E-05	0,0004371	1,67E-06	0,000132	4,35E-05	1,84E-05	0,001041	0	3,35E-06	0	3,35E-06	0	2,68E-05	0,000136	5,02E-06	2,18E-05	0,00112	
д	0,007259	0,000211	8,37E-06	0,000278	0,001292	0,001706	0,002501	0,000571	0,000773	0,01443	0,00172	2,68E-05	7,70E-05	0,000143	0,004594	1,67E-06	0,000347	0,000305	9,34E-05	0,001805	0	0	0	0	0	0	6,86E-05	0,000126	0,001517	0,000263	0,00065		
е	0,003651	0,001185	0,001955	0,005141	0,000393	0,004557	0,000411	0,003599	0,000392	0,001818	3,35E-06	0,000735	0,004157	0,00336	0,009043	0,001964	0,002066	0,002244	0,002843	0,000421	0,000122	0,000104	6,86E-05	0,000738	0,003415	0,00191	0,001482	0,000408	0,001085	0	1,67E-06	9,71E-05	
ж	0,001671	0,001152	2,68E-05	3,35E-06	0	0,000706	0,000757	1,67E-06	0,000144	0,000365	6,70E-06	0,000259	0	1,67E-05	0,002138	0	0,000266	4,69E-05	5,02E-06	0,001152	0	0	0	0	0	0	0	0	4,85E-05	0	3,52E-05	0,00014	
з	0,00459	0,003554	1,51E-05	0,000549	3,35E-06	1,17E-05	0,001354	0	8,87E-05	0,002071	1,51E-05	4,69E-05	3,35E-06	9,21E-05	0,001334	0	3,18E-05	1,17E-05	1,00E-05	0,000233	0	0	0	0	0	0	0	0	0,000137	7,87E-05	0	0,000276	0,00023
и	0,008402	0,000109	0,000727	0,002804	0,000588	0,003835	0,00121	0,001428	0,000465	0,000425	1,67E-06	0,002784	0,006444	0,002833	0,000698	0,00036	0,000882	0,003879	0,001455	0,001199	3,85E-05	0,000142	0,000249	0,000263	0,001796	0,001435	0,001003	4,52E-05	1,51E-05	0	1,67E-06	5,02E-06	
й	5,02E-06	0,005554	0	0	0	0	0,002042	0	0	0,001038	0	0	0	0	0,003814	0	0	0	0	0	0,000121	0	0	0	0	0	0	0	0,001515	7,87E-05	6,33E-05	3,18E-05	
к	0,008906	0,004261	0,000157	0,000208	0,000134	0,000241	0,001659	0,000102	0,000203	0,00199	0,000368	4,69E-05	0,00075	0,000169	0,000355	0,003408	0,000134	0,000408	0,002998	0,000752	0,001115	0	1,67E-06	1,67E-05	0,000509	0,000583	0	0,000932	0,000164	6,03E-05	9,21E-05	0,00021	
л	0,002431	0,008946	0,000691	0,000551	0,001647	0,000514	0,000153	8,37E-06	0,000265	0,004898	3,68E-05	0,000732	0,000229	0,000239	3,35E-06	0,000728	0,000815	0,000147	0,002821	0,000186	4,85E-05	0,000157	0	8,37E-06	0,000422	0	0	0,000233	8,04E-05	5,02E-06	0,00095		
м	0,006321	0,002712	4,02E-05	0,000104	6,70E-06	6,53E-05	0,003599	1,67E-06	0,000425	0,002384	0,000104	8,37E-06	7,73E-05	4,52E-05	3,35E-06	0,005166	0	0,000246	0,000964	4,69E-05	0,001436	3,35E-06	6,19E-05	0	0	0	0	0,000196	0,001393	2,18E-05	6,70E-05	0,00036	
н	0,015196	0,004063	0,000353	0,001289	0,00041	0,00188	0,006141	0,000939	0,001552	0,003362	0,000398	0,000425	0,000475	0,001209	0,00291	0,006757	0,000213	0,001266	0,000996	0,001021	0,000419	0	0,000233	0	0,000922	0,000413	8,04E-05	0,001299	0,000162	0,000131	8,54E-05	0,00073	
о	0	5,69E-05	0,00198	0,005673	0,006961	0,003259	0,001741	0,000239	0,000512	0,000184	4,52E-05	0,007917	0,007438	0,003648	0,01009	0,000208	0,009598	0,007974	0,002302	0,0127	0,000104	0,000179	0,001932	0,000256	0,000112	0,00028	3,35E-06	3,18E-05	5,02E-06	0	3,52E-06	6,70E-06	
п	0,015859	0,000894	0	0,00031	1,67E-06	7,20E-05	0,000076	0	1,51E-05	0,000254	2,01E-05	6,70E-06	4,02E-05	0,000107	0,000143	2,18E-05	0,000116	0,001687	0,000107	0,000663	0	3,35E-06	1,67E-06	0	1,51E-05	0	0	1,17E-05	0,000219	6,70E-06	3,00E-06	4,19E-05	
р	0,005252	0,002222	0,0012	0,000713	0,0008	0,001055	0,006567	1,84E-05	0,0003	0,000954	7,73E-05	0,001848	8,37E-06	7,73E-05	0,000288	0,000504	0,005997	3,01E-05	0,000519	0,0004415	0,000656	2,68E-05	0,000167	1,67E-06	1,17E-05	2,34E-05	6,70E-06	1,17E-05	0,000325	4,02E-05	2,18E-05	0,00012	
с	0,016028	0,003365	0,000913	0,002118	1,34E-05	0,000186	0,003658	5,02E-06	3,35E-06	0,002568	0,000194	0,001771	0,000146	0,00131	0,005806	3,01E-05	0,000211	0,000579	0,001199	0,001378	5,02E-06	4,85E-05	0	3,35E-06	0	0	0	0,000896	0,000685	1,00E-05	0,000144	0,00063	
т	0,008193	0,004048	3,35E-06	0,000244	2,01E-05	0,000159	0,003502	3,35E-06	1,84E-05	0,003082	0,000328	0,000446	6,03E-05	0,000537	0,006762	0,000137	0,001081	0,010682	0,000519	0,001664	6,70E-06	2,68E-05	0	0,000345	0,000147	0	0	8,87E-05	0,000743	0,002371	0,000313	0,00138	
у	0,004446	0,000194	0,00112	0,000732	0,000467	0,001639	8,87E-05	0,000221	0,000388	8,37E-06	1,67E-06	0,000222	0,001122	0,001823	0,00349	8,87E-05	0,000966	0,002625	0,000737	0,000143	1,67E-05	0,000107	0,000137	0,000131	0,000794	0,000228	0,000218	1,67E-06	5,02E-06	0	0	1,67E-06	
ф	0,000442	3,68E-05	0	0	0	0	4,85E-05	0	0	2,85E-05	1,34E-05	0	8,37E-06	1,67E-06	1,84E-05	7,73E-05	0	1,84E-05	2,18E-05	1,17E-05	1,34E-05	1,34E-05	0	0	0	0	0	5,02E-06	1,67E-06	3,35E-06	0	0	
х	0,001364	0,000855	5,69E-05	3,01E-05	1,67E-06	5,02E-05	0,00041	0	0	0,001393	0	1,67E-06	3,35E-06	6,70E-06	5,02E-06	0,000628	0	0,000196	0,000204	1,34E-05	0,000489	0	6,70E-06	0	3,35E-06	0	0	3,35E-06	0,000825	2,01E-05	3,68E-05	0,00011	
ц	0,000253	0,000164	6,70E-06	3,01E-05	0	0,000208	0,000291	0	1,67E-06	0,00075	2,68E-05	0	5,02E-06	0,000285	8,87E-05	5,02E-05	5,02E-05	4,35E-05	6,53E-05	1,17E-05	0	1,67E-06	8,37E-06	0	1,17E-05	0	0	0,000451	1,17E-05	0	5,02E-06	2,01E-05	
ч	0,005494	0,000679	1,67E-06	3,35E-05	3,18E-05	1,67E-05	0,000839	7,70E-05	3,35E-06	0,000862	1,51E-05	5,02E-06	0,000105	5,86E-05	0,000208	0,001878	2,68E-05	0,000157	0,000281	0,00114	0,0008	0	0	5,02E-06	0	0	0	0,000454	0,000169	3,01E-05	0,000152	0,00015	
ш	0,000936	0,000711	5,02E-06	0,000879	0	3,68E-05	0,000929	0	3,35E-06	0,000722	5,86E-05	3,01E-05	2,34E-05	3,35E-06	1,67E-06	0,000857	5,02E-06	0,000162	7,87E-05	2,34E-05	0,000494	0	1,51E-05	0	8,87E-05	6,70E-06	0	0,000568	0,000487	5,02E-06	4,69E-05	1,84E-05	
щ	0,000126	0,00034	9,04E-05	1,34E-05	0	0,000691	0	0	0,000186	0	1,67E-06	8,37E-06	1,67E-06	0,000199	0,000253	0	3,01E-05	1,67E-06	1,34E-05	0,000305	0	0	0	0	0	0	0	5,02E-06	2,34E-05	0	0,000581	0,00029	
ь	0	0	0,000174	0,000285	0	0,000529	0	4,69E-05	0,000162	0	0	0	0,000452	5,86E-05	0,00102	0	4,35E-05	0,000713	0,003082	0,000589	0	1,67E-06	0	0,000239	0,000872	1,34E-05	0	0	0	0	0	0	
ы	0	0	0,003713	0,0022	0	0,000685	0	0	0,000362	0	0	0	0,000889	0,000782	0,003249	0	0,000385	0,0001458	0,000315	0,001903	0	1,34E-05	0	0,000196	0	0	0	0	0	0	0	0	
э	0,004415	2,18E-05	3,35E-06	5,52E-05	0	4,02E-05	6,70E-06	0	1,17E-05	1,67E-06	1,34E-05	1,34E-05	2,51E-05	1,67E-06	6,19E-05	0,00736	1,00E-05	1,17E-05	6,70E-06	2,01E-05	0	5,86E-05	0	3,35E-06	3,35E-06	0	8,37E-06	0	5,02E-06	3,35E-06	6,70E-06		
ю	0,000121	0,000994	3,35E-06	0	0	9,71E-05	0,000109	0	0	0,000171	0	1,67E-06	0,000768	0	0,000238	0,000166	6,70E-06	0,000152	0,000363	5,52E-05	0,00109	0	0	0	0	0	0	0	0,000491	0	2,68E-05	9,71E-05	
я	0,001505	0,001862	0,000435	0,000298	0	0,000258	0,000161	0	0,000154	0,000897	0	1,67E-06	0,001744	0,000502	0,001299	0,000633	0,000395	0,000949	0,004267	0,0000491	8,78E-05	1,67E-06	0	0	0	0	0	0	0,00036	1,84E-05	1,67E-06	0	8,71E-05

Без пробілів:

H1 і H2:

```
*****h1*****  
4.467371729797125  
  
*****h2*****  
4.155390354697318  
  
*****h2_nocross*****  
4.155210790886074
```

Частоти букв:

---

а: 0.07867052069181213  
б: 0.017015414772523538  
в: 0.041644587697664874  
г: 0.017527381241707937  
д: 0.03489539548911723  
е: 0.08232799321096604  
ж: 0.010536150410569219  
з: 0.01771662954743758  
и: 0.062857330682529  
й: 0.010994330519177826  
к: 0.03656277266696681  
л: 0.05455231819214082  
м: 0.03044009195475571  
н: 0.06562633220846796  
о: 0.11335774304462674  
п: 0.027255740199925896  
р: 0.045399672500826715  
с: 0.05051734510524198  
т: 0.06161924834557139  
у: 0.030114385660157852  
ф: 0.0009263206543608786  
х: 0.00796536158378919  
ц: 0.003083751339678796  
ч: 0.01556816325554896  
ш: 0.008656615921559566  
щ: 0.0037969186391652358  
ы: 0.01918778611355695  
ь: 0.020179845653066023  
э: 0.0053557270521488645  
ю: 0.005882634177048762  
я: 0.019765491467889543



## Частоти біграм з пересіченнями без пробілів:

	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	ь	ы	э	ю	я	
а	0,000274	0,001205	0,005806	0,001285	0,005099	0,000229	0,001446	0,006808	0,000308	0,000104	0,008784	0,008272	0,00371325	0,011341	0,000154	0,002	0,007985	0,001467	0,006908	0,000212	0,000143	0,000508	0,000789	0,001986	0,001002	0,000478	0,000135	5,28E-05	0	5,58E-05	0,00012	
б	0,001597	1,29E-05	0,0002	7,07E-05	0,000106	0,002478	2,49E-05	0,000206	0,001534	0,000386	0,000283	0,000364	0,00034463	0,00022	0,000338	0	9,26E-05	0,000182	0,000469	0,0001209	6,97E-06	0,000121	1,49E-05	9,96E-07	7,97E-06	0	0,000508	0,000526	1,99E-06	0,00029	0,00046	
в	0,004871	6,47E-05	0,000379	8,77E-05	0,001642	0,003704	2,49E-05	0,001129	0,000446	0,000707	0,000283	0,000822	0,00075699	0,000457	0,000193	9,96E-06	0,00067	0,001628	0,000469	0,0001682	9,96E-07	0,00058	0,0000114	2,19E-05	4,18E-05	1,99E-06	0,000481	0,0001583	8,96E-06	0,000202	0,001553	
г	0,001278	9,96E-06	0,000341	1,79E-05	0,000201	0,004259	2,29E-05	0,000551	0,001054	0,000267	0,00014	0,00047	0,00028885	0,000155	0,000591	3,98E-06	0,000177	0,000167	0,000145	0,0001361	0	8,77E-05	4,98E-06	2,99E-06	0	0	0,000205	0,000253	8,96E-06	7,97E-05	0,000383	
д	0,003237	1,89E-05	0,000636	0,001579	0,000104	0,003797	0,000729	0,001056	0,002788	0,000061	0,000341	0,00047	0,00043228	0,000611	0,000581	5,98E-06	0,000487	0,000511	0,000407	0,0002375	2,99E-06	0,000167	2,09E-05	9,96E-07	2,99E-06	1,99E-06	0,000632	0,000367	0,001942	0,00046	0,001293	
е	0,001843	0,002388	0,006257	0,000458	0,005492	0,001901	0,004326	0,000505	0,002593	0,00011	0,000957	0,005309	0,00426705	0,010901	0,002869	0,002426	0,000867	0,003475	0,005345	0,00024	0,000136	0,000134	0,00085	0,004128	0,002292	0,001723	0,000902	0,001441	0	4,68E-05	0,000347	
ж	0,001526	2,29E-05	7,97E-05	4,98E-06	0,000892	0,001086	1,89E-05	0,000146	0,000576	0,000154	0,000135	0,000363	8,068E-05	4,78E-05	0,002759	0	0,000357	9,76E-05	7,57E-05	0,0001492	0	2,19E-05	1,99E-06	0	9,96E-07	0	7,87E-05	0,000101	0	0,000104	0,000314	
з	0,004839	1,89E-05	0,000824	4,88E-05	9,96E-05	0,002357	1,2E-05	0,000146	0,003085	0,000254	0,000176	0,000272	0,00019722	0,000286	0,002358	3,98E-06	7,27E-05	8,37E-05	0,000155	0,000462	0	8,47E-05	9,96E-06	2,99E-06	9,96E-07	0	0,000481	0,000311	1,99E-06	0,000408	0,000671	
и	0,001205	0,000892	0,003551	0,000743	0,000462	0,000975	0,001648	0,000599	0,001616	0,000447	0,003781	0,008229	0,00393039	0,008289	0,0001898	0,00103	0,004736	0,0001825	0,004126	0,000534	0,000157	0,000452	0,000337	0,002196	0,001661	0,001194	0,000848	0,000386	0	0,000174	0,000732	
й	0,000674	0	0	0	0,002433	0	0	0,00124	9,96E-07	0	0	0	0	0	0,004514	0	0	0	0	0,000164	0	0	0	0	0	0	0	0,001766	8,17E-05	7,47E-05	4,48E-05	
к	0,006129	0,000184	0,000799	0,000195	0,000441	0,00289	0,000133	0,000378	0,003485	0,000997	0,000403	0,001477	0,0007052	0,000573	0,005237	0,000166	0,000544	0,003825	0,001196	0,00175	9,96E-07	0,000236	3,88E-05	0,000598	0,000688	0	0,001809	0,000416	8,37E-05	0,000296	0,000886	
л	0,011084	0,000831	0,000815	0,002033	0,000689	0,00878	1,1E-05	0,00034	0,006073	0,000188	0,001001	0,000354	0,00045021	4,98E-05	0,009066	0,000981	0,000198	0,00337	0,000359	0,002688	6,47E-05	0,000302	5,98E-06	1,49E-05	0,000475	0	0,000229	0,002712	8,37E-05	6,67E-05	0,001278	
м	0,004011	5,88E-05	0,000419	3,19E-05	0,000118	0,005088	2,19E-05	0,000592	0,003509	0,000538	0,00032	0,000396	0,00041037	0,000138	0,006953	2,99E-05	0,000321	0,001322	0,000336	0,002032	5,98E-06	0,000211	7,97E-06	5,98E-06	4,18E-05	9,96E-07	0,000755	0,001799	2,59E-05	0,00021	0,000758	
н	0,006902	0,000462	0,002029	0,000565	0,002521	0,008993	0,00121	0,001979	0,005941	0,000957	0,001123	0,001404	0,00220823	0,003854	0,010276	0,000264	0,00161	0,0015	0,001807	0,001534	1,99E-06	0,000541	4,18E-05	0,001097	0,000539	8,57E-05	0,002739	0,000731	0,000162	0,000328	0,002221	
о	0,000152	0,002379	0,007159	0,008438	0,004016	0,001569	0,000324	0,000754	0,001899	0,00049	0,009986	0,009724	0,00504695	0,012213	0,002092	0,011221	0,009501	0,02914	0,015682	0,000543	0,000202	0,002567	0,000338	0,000154	0,000316	2,99E-06	0,000914	0,000432	0	0,000201	0,000757	
п	0,003108	1,79E-05	0,00103	0,000104	0,000404	0,003169	4,08E-05	0,000231	0,002691	0,000849	0,000714	0,000711	0,00094823	0,000621	0,004065	3,29E-05	0,000244	0,002312	0,000657	0,001408	1,99E-06	0,000432	3,39E-05	1,29E-05	2,39E-05	9,96E-07	0,001101	0,000784	9,96E-06	0,000321	0,001176	
р	0,00328	0,001411	0,00108	0,001023	0,001349	0,008351	2,89E-05	0,000418	0,001854	0,000573	0,002491	0,000406	0,00036356	0,000444	0,006732	7,27E-05	0,000299	0,005508	0,000975	3,88E-05	0,000335	1,69E-05	2,09E-05	2,99E-05	6,97E-06	0,000329	0,000559	3,39E-05	0,000155	0,000494		
с	0,005952	0,000225	0,00321	0,000107	0,000417	0,006417	3,88E-05	0,000184	0,005131	0,001453	0,00094	0,0003106	0,00102891	0,000853	0,009455	4,58E-05	0,000394	0,000932	0,00199	0,002256	3,98E-06	0,000391	3,39E-05	1,39E-05	9,96E-06	0,0002258	0,0001283	1,39E-05	0,000454	0,00189		
т	0,005821	1,49E-05	0,000731	6,47E-05	0,000351	0,007403	2,99E-05	0,000161	0,004769	0,000652	0,000873	0,000531	0,00040439	0,000809	0,00932	0,000146	0,001385	0,012981	0,000955	0,002337	8,96E-06	0,000175	9,96E-06	0,00398	0,000174	0	0,000765	0,001181	0,002829	0,000559	0,002197	
у	0,000773	0,001322	0,00092	0,000582	0,002084	0,00073	0,000288	0,000522	0,000694	0,000234	0,002573	0,001507	0,00242537	0,004424	0,000781	0,001163	0,003214	0,000992	0,001819	0,000162	0,000119	0,000269	0,000163	0,000922	0,000276	0,000271	0,000308	0,000231	0	5,38E-05	0,000302	
ф	0,000108	9,96E-07	2,79E-05	9,96E-07	1,99E-06	9,96E-05	0	1,2E-05	8,17E-05	7,07E-05	1,49E-05	2,59E-05	2,1913E-05	2,29E-05	0,000151	0	3,09E-05	4,88E-05	4,08E-05	3,29E-05	1,1E-05	2,79E-05	0	0	9,96E-07	9,96E-07	3,98E-05	8,96E-06	5,98E-06	1,39E-05	2,69E-05	
х	0,001216	5,78E-05	7,07E-05	2,99E-06	8,27E-05	0,000725	3,98E-06	9,96E-06	0,001825	8,47E-05	4,68E-05	5,88E-05	7,3707E-05	4,68E-05	0,000925	0	0,00024	0,000274	5,88E-05	0,000645	0	2,49E-05	2,99E-06	0	0	0,000106	0,001057	1,59E-05	5,78E-05	0,000248		
ц	0,000237	7,97E-06	8,17E-05	1,99E-06	0,000262	0,000361	0	1,2E-05	0,000919	4,78E-05	4,68E-05	3,98E-06	2,5897E-05	0,000361	0,000144	6,97E-06	6,08E-05	0,000102	2,29E-05	0	9,96E-06	1,2E-05	0	1,1E-05	0	0	0,000217	2,19E-05	0	9,96E-06	3,98E-05	
ч	0,001588	1,2E-05	0,00039	4,38E-05	8,07E-05	0,001598	8,96E-05	5,28E-05	0,001487	0,000453	0,00021	0,000813	0,00038248	0,000326	0,003283	3,29E-05	0,000188	0,000397	0,000437	0,0001287	2,99E-06	8,96E-05	6,97E-06	5,98E-06	0	0	0,001093	0,000328	2,89E-05	0,000326	0,000536	
ш	0,000966	1,2E-05	0,001109	1,2E-05	8,27E-05	0,00126	1,99E-06	1,99E-05	0,000972	0,000171	6,77E-05	5,279E-05	3,39E-05	0,001166	0,000203	0,000155	5,68E-05	0,000621	0,001819	0,000162	0,000119	0,000269	0,000163	0,000922	0,000276	0,000271	0,000308	0,000231	7,97E-06	7,17E-05	9,96E-05	
щ	0,000423	0,000111	1,89E-05	0	3,98E-06	0,000842	0	3,98E-06	0,000235	1,99E-05	1,2E-05	1,39E-05	6,9723E-06	0,000258	0,000332	0	2,79E-05	3,98E-06	1,79E-05	0,000374	0	3,98E-06	0	0	0	0	0	1,39E-05	2,09E-05	0	0,000696	0,000358
ь	0	0,000209	0,000336	0	0,000578	0	4,78E-05	0,000199	0	0	9,96E-07	0,004976	6,3747E-05	0,001173	0	4,38E-05	0,000811	0,003662	0,006734	0	9,96E-07	1,99E-06	0	0,000278	0,00104	2,39E-05	0	0	0	0	0	0
ы	0	0,000455	0,002657	2,99E-06	0,00076	0	0	0,00044	0	0	0	0,001033	0,0009064	0,003895	0	0,00047	0,001743	0,000402	0,002193	0	1,2E-05	0	0,000217	0	0	0	0	0	0	0	0	0
э	0,000621	5,78E-05	0,000326	1,69E-05	8,27E-05	0,000379	9,96E-06	5,08E-05	0,000513	0,000113	0,0002	0,000475	0,00018327	6,18E-05	0,00071	3,98E-06	4,38E-05	0,000127	0,002022	0,000211	9,96E-07	0,000122	1,1E-05	6,97E-06	9,96E-06	0	0,000348	0,000118	3,98E-06	7,17E-05	0,000256	
ю	0,00119	3,98E-06	5,98E-06	0	0,000122	0	9,96E-07	0,000235	4,98E-06	3,98E-06	0,000894	9,9604E-06	0,000249	0,00021	1,2E-05	0,000192	0,000337	7,07E-05	0,000136	0	0	0	0	0	0	0	0,000628	2,99E-06	0	4,78E-05	0,000124	
я	0,002459	0,000548	0,000356	7,97E-06	0,000317	0,000338	2,99E-06	0,000209	0,001304	6,77E-05	7,07E-05	0,002007	0,00071018	0,00165	0,001072	0,000472	0,00113	0,005118	0,000832	0,000168	9,96E-07	2,19E-05	0	0	1,99E-06	0	0,000561	7,87E-05	2,99E-06	4,78E-05	0,000209	

## Частоти біграм без пересічень без пробілів:

	а	б	в	г</
--	---	---	---	-----

2. За допомогою програми CoolPinkProgram оцінити значення (10) H , (20) H , (30) H .

H(10)

Произвольная часть текста:  
нали\_зара

Использованные буквы:  
ж,

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

Введенный символ: ж

Символ по счету: 1

Номер эксперимента: 54

Поле ввода символов:

Неравенство для энтропии:  
 $1.22679317516447 < H < 2.01415566370604$

Двоичная таблица угаданных символов:

Вероятности:

- q[1] = 0.6226415
- q[2] = 0.0754716
- q[3] = 0.0943396
- q[4] = 0.0566037
- q[5] = 0.0377358
- q[6] = 0.0377358
- q[7] = 0.0188679
- q[8] = 0.0377358
- q[9] = 0
- q[10] = 0
- q[11] = 0
- q[12] = 0
- q[13] = 0
- q[14] = 0
- q[15] = 0
- q[16] = 0
- q[17] = 0
- q[18] = 0
- q[19] = 0.018867
- q[20] = 0
- q[21] = 0
- q[22] = 0
- q[23] = 0
- q[24] = 0
- q[25] = 0
- q[26] = 0
- q[27] = 0
- q[28] = 0
- q[29] = 0
- q[30] = 0
- q[31] = 0
- q[32] = 0

Строка состояния:  
Вы не угадали. Введите другую букву

H(20)

Произвольная часть текста:  
мной\_в\_этом\_пункте\_

Использованные буквы:

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

Введенный символ:

Символ по счету:

Номер эксперимента: 51

Поле ввода символов:

Неравенство для энтропии:  
 $1.62679328647681 < H < 2.34153709471428$

Двоичная таблица угаданных символов:

Вероятности:

- q[1] = 0.5
- q[2] = 0.24
- q[3] = 0.04
- q[4] = 0.04
- q[5] = 0.02
- q[6] = 0
- q[7] = 0.04
- q[8] = 0
- q[9] = 0
- q[10] = 0
- q[11] = 0.02
- q[12] = 0
- q[13] = 0.02
- q[14] = 0.02
- q[15] = 0
- q[16] = 0
- q[17] = 0
- q[18] = 0
- q[19] = 0
- q[20] = 0.02
- q[21] = 0
- q[22] = 0
- q[23] = 0
- q[24] = 0
- q[25] = 0
- q[26] = 0
- q[27] = 0
- q[28] = 0
- q[29] = 0.02
- q[30] = 0.02
- q[31] = 0
- q[32] = 0

Строка состояния:

H(30)

3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

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

$H_0 = \log_2 32 = 5$  (32 це довжина алфавіта)

з пробілами

$H_1$

$$R = 1 - \frac{4.387662336862179}{5} = 0,122467532627$$

$H_2$

$$R = 1 - \frac{3.9909063611403606}{5} = 0,201818727771$$

$H_2$  (nocross)

$$R = 1 - \frac{3.989031146917921}{5} = 0,2021937706164158$$

без пробілів

$H_1$

$$R = 1 - \frac{4.467371729797125}{5} = 0,106525654040575$$

$H_2$

$$R = 1 - \frac{4.155390354697318}{5} = 0,1689219290605364$$

$H_2$  (nocross)

$$R = 1 - \frac{4.155210790886074}{5} = 0,1689578418227852$$



H(10)

$$R_L = 1 - \frac{1.22679317516447}{5} = 0,754641364967$$

$$R_R = 1 - \frac{2.01415566370604}{5} = 0,597168867258$$

H(20)

$$R_L = 1 - \frac{1.62679328647681}{5} = 0,674641342704$$

$$R_R = 1 - \frac{2.34153709471428}{5} = 0,531692581057$$

H(30)

$$R_L = 1 - \frac{0.647316791063824}{5} = 0,870536641787$$

$$R_R = 1 - \frac{1.2323876647553}{5} = 0,753522467048$$

Висновки: вирахувавши частоту появи символів та груп символів у довільному тексті, можна визначити їх ентропію, яка є мірою кількості інформації. Коли джерело даних має менш імовірне значення (наприклад, коли стається низькоймовірна подія), то ця подія несе більше «інформації» («неочікуваності»), ніж коли джерело даних має більш імовірне значення. Визначена таким чином кількість інформації, що передається кожною подією, стає випадковою змінною, чие математичне сподівання є інформаційною ентропією. За отриманими результатами можна вирахувати приблизне значення надлишковості мови, тобто наскільки легко вгадати наступний символ, знаючи n попередніх.

„освіська мова - лайно, вивчайте українську“