

Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Фізико-технічний інститут

СИМЕТРИЧНА КРИПТОГРАФІЯ

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Експериментальна оцінка ентропії на символ джерела відкритого тексту

Виконали:

студентки групи ФІ-94

Зацаренко А. Ю.

Футурська О.В.

Перевірив:

Чорний О.М.

ЗМІСТ

ЗАГАЛЬНІ ВІДОМОСТІ	3
1. Мета комп'ютерного практикуму	3
2. Постановка задачі	3
3. Хід роботи	3
4. Опис труднощів	3
ПРАКТИЧНА ЧАСТИНА	4
1. Результати підрахунку частот	4
3. Значення ентропії	10
4. Оцінка надлишковості російської мови	12
ВИСНОВКИ	13

ЗАГАЛЬНІ ВІДОМОСТІ

1. Мета комп'ютерного практикуму

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.

2. Постановка задачі

Створити програму для експериментальної оцінки ентропії на символ джерела відкритого тексту, порівняти різні моделі джерела відкритого тексту для наближеного визначення ентропії.

3. Хід роботи

1. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
2. 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 видалено всі пробіли.
3. За допомогою програми CoolPinkProgram оцінити значення $H^{(10)}$, $H^{(20)}$, $H^{(30)}$.
4. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

4. Опис труднощів

Реалізуючи програмний код, ми зіштовхнулися з проблемою зайвих символів у тексті при зчитуванні літер на межі строк. В ASCII дані символи розташовані на 10 місці (LINE FEED) та на 13 (CARRIAGE RETURN). В результаті ми вирішили вчинити таким чином: видаляти символи, а на їх місцях записувати пробіли. На коректність програми це не впливає, адже потім ми використовуємо функцію `Regex.Replace` з бібліотеки `System.Text.RegularExpressions`, яка замінює усі пробіли одним.

ПРАКТИЧНА ЧАСТИНА

1. Результати підрахунку частот

а	0,06687	д	0,02593	и	0,05466	м	0,02503	р	0,03436	ф	0,00067	ш	0,00864	э	0,00185
б	0,01494	е	0,07130	й	0,00818	н	0,05227	с	0,04373	х	0,00781	щ	0,00224	ю	0,00572
в	0,03841	ж	0,00866	к	0,03378	о	0,09296	т	0,05062	ц	0,00386	ы	0,01595	я	0,01689
г	0,01613	з	0,01412	л	0,03959	п	0,02234	у	0,02616	ч	0,01369	ь	0,01700		0,16565

Табл.2.1 – Частоти букв в алфавіті з пробілом

а	0,08015	д	0,03108	и	0,06552	м	0,03000	р	0,04118	ф	0,00080	ш	0,01035	э	0,00221
б	0,01790	е	0,08546	й	0,00981	н	0,06264	с	0,05241	х	0,00936	щ	0,00269	ю	0,00685
в	0,04604	ж	0,01038	к	0,04049	о	0,11141	т	0,06067	ц	0,00462	ы	0,01912	я	0,02024
г	0,01933	з	0,01692	л	0,04745	п	0,02677	у	0,03136	ч	0,01641	ь	0,02037		

Табл.2.2 – Частоти букв в алфавіті без пробілу

аб	0,00071	ав	0,00288	аг	0,00054	ад	0,00164	ае	0,00103	аж	0,00111	аз	0,00341	аи	0,00008
ай	0,00074	ак	0,00668	ал	0,00745	ам	0,00341	ан	0,00353	ао	0,00001	ап	0,00088	ар	0,00277
ас	0,00334	ат	0,00461	ау	0,00009	аф	0,00006	ах	0,00095	ац	0,00015	ач	0,00087	аш	0,00088
ащ	0,00026	аэ	0,00002	аю	0,00071	ая	0,00190	а	0,01624	ба	0,00107	бб	0,00001	бв	0,00002
бг	0,00001	бд	0,00001	бе	0,00208	бж	0,00003	бз	0,00001	би	0,00054	бк	0,00018	бл	0,00083
бм	0,00004	бн	0,00016	бо	0,00203	бр	0,00136	бс	0,00011	бт	0,00001	бу	0,00147	бх	0,00002
бц	0,00002	бч	0,00002	бш	0,00002	бщ	0,00008	бы	0,00378	бь	0,00012	бю	0,00000	бя	0,00051
б	0,00036	ва	0,00567	вб	0,00001	вв	0,00008	вг	0,00001	вд	0,00032	ве	0,00499	вж	0,00000
вз	0,00041	ви	0,00371	вк	0,00022	вл	0,00062	vm	0,00025	вн	0,00089	во	0,00654	вп	0,00023
вр	0,00050	вс	0,00320	вт	0,00007	ву	0,00067	vx	0,00002	вц	0,00002	вч	0,00008	вш	0,00146
вщ	0,00005	вы	0,00249	вь	0,00017	во	0,00000	вя	0,00029	в	0,00564	га	0,00092	гб	0,00001
гв	0,00002	гд	0,00119	ге	0,00026	ги	0,00054	гк	0,00011	гл	0,00146	гм	0,00001	гн	0,00030
го	0,00875	гр	0,00092	гс	0,00002	гт	0,00002	гу	0,00066	гч	0,00001	гш	0,00001	г	0,00076
да	0,00432	дб	0,00006	дв	0,00080	дг	0,00002	дд	0,00001	де	0,00453	дж	0,00003	дз	0,00001
ди	0,00223	дк	0,00031	дл	0,00049	дм	0,00006	дн	0,00215	до	0,00358	дп	0,00006	др	0,00112
дс	0,00022	дт	0,00042	ду	0,00192	дх	0,00004	дц	0,00025	дч	0,00002	дш	0,00009	ды	0,00072
дь	0,00083	дю	0,00002	дя	0,00032	д	0,00140	еа	0,00003	еб	0,00144	ев	0,00197	ег	0,00350
ед	0,00267	ее	0,00142	еж	0,00089	ез	0,00124	еи	0,00011	ей	0,00174	ек	0,00111	ел	0,00551
ем	0,00387	ен	0,00558	ео	0,00016	еп	0,00104	ер	0,00548	ес	0,00450	ет	0,00564	еу	0,00007
еф	0,00000	ех	0,00065	ец	0,00097	еч	0,00111	еш	0,00095	ещ	0,00080	ею	0,00044	ея	0,00024
е	0,01823	жа	0,00111	жб	0,00002	жг	0,00001	жд	0,00058	же	0,00351	жж	0,00002	жи	0,00135
жк	0,00014	жл	0,00002	жм	0,00001	жн	0,00089	жо	0,00004	жр	0,00001	жс	0,00004	жу	0,00022
жц	0,00006	жч	0,00001	жь	0,00005	ж	0,00043	за	0,00548	зб	0,00013	зв	0,00081	зг	0,00023
зд	0,00068	зе	0,00047	зж	0,00008	зз	0,00000	зи	0,00039	зк	0,00012	зл	0,00022	зм	0,00019
зн	0,00166	зо	0,00063	зр	0,00014	зс	0,00001	зт	0,00002	зу	0,00026	зч	0,00006	зш	0,00001
зы	0,00052	зь	0,00018	зю	0,00000	зя	0,00030	з	0,00136	иа	0,00006	иб	0,00058	ив	0,00251
иг	0,00046	ид	0,00162	ие	0,00179	иж	0,00033	из	0,00202	ии	0,00028	ий	0,00115	ик	0,00219
ил	0,00470	им	0,00232	ин	0,00307	ио	0,00007	ип	0,00014	ир	0,00072	ис	0,00270	ит	0,00364
иу	0,00001	иф	0,00004	их	0,00168	иц	0,00112	ич	0,00151	иш	0,00048	ищ	0,00014	ию	0,00021
ия	0,00076	и	0,01819	йв	0,00000	йд	0,00018	йз	0,00001	йк	0,00010	йл	0,00000	йм	0,00003
йн	0,00025	йо	0,00006	йр	0,00000	йс	0,00030	йт	0,00018	йц	0,00001	йч	0,00006	йш	0,00008
йя	0,00000	й	0,00699	ка	0,00882	кв	0,00028	кд	0,00001	ке	0,00053	кж	0,00003	кз	0,00000
ки	0,00366	кк	0,00001	кл	0,00053	км	0,00001	кн	0,00056	ко	0,00904	кр	0,00172	кс	0,00008
кт	0,00046	ку	0,00239	кц	0,00001	кш	0,00000	к	0,00538	ла	0,00538	лб	0,00005	лв	0,00004
лг	0,00018	лд	0,00011	ле	0,00386	лж	0,00025	лз	0,00001	ли	0,00586	лк	0,00034	лл	0,00014
лм	0,00000	лн	0,00028	ло	0,00637	лп	0,00009	лс	0,00154	лт	0,00007	лу	0,00138	лч	0,00011
лш	0,00001	лщ	0,00001	лы	0,00067	ль	0,00369	лю	0,00087	ля	0,00168	л	0,00673	ма	0,00231
мб	0,00002	мг	0,00003	ме	0,00363	ми	0,00295	мк	0,00015	мл	0,00030	мм	0,00002	мн	0,00148
мо	0,00308	мп	0,00005	мр	0,00004	мс	0,00009	мт	0,00001	му	0,00217	мф	0,00001	мц	0,00004
мч	0,00004	мщ	0,00000	мы	0,00079	мь	0,00005	мя	0,00039	м	0,00767	на	0,00959	нб	0,00003
нв	0,00000	нг	0,00004	нд	0,00023	не	0,01034	нж	0,00000	нз	0,00002	ни	0,00639	нк	0,00029
нм	0,00000	нн	0,00253	но	0,00924	нр	0,00003	нс	0,00024	нт	0,00026	ну	0,00274	нф	0,00000
нц	0,00017	нч	0,00011	нш	0,00000	нщ	0,00004	ны	0,00329	нь	0,00143	ню	0,00019	ня	0,00129
н	0,00380	об	0,00321	ов	0,00695	ог	0,00495	од	0,00482	ое	0,00207	ож	0,00190	оз	0,00138

ои	0,00098	ой	0,00258	ок	0,00219	ол	0,00612	ом	0,00477	он	0,00513	оо	0,00012	оп	0,00124
ор	0,00578	ос	0,00535	от	0,00592	оу	0,00005	оф	0,00015	ох	0,00045	оц	0,00007	оч	0,00181
ош	0,00089	ощ	0,00017	оз	0,00002	ою	0,00135	оя	0,00059	о	0,02233	па	0,00164	пе	0,00243
пи	0,00105	пк	0,00022	пл	0,00066	пн	0,00009	по	0,00866	пп	0,00001	пр	0,00616	пс	0,00002
пт	0,00005	пу	0,00066	пц	0,00006	пч	0,00004	пш	0,00002	пщ	0,00001	пы	0,00022	пъ	0,00008
пя	0,00026	п	0,00009	ра	0,00623	рб	0,00007	рв	0,00020	рг	0,00015	рд	0,00030	ре	0,00543
рж	0,00022	рз	0,00004	ри	0,00479	рк	0,00041	рл	0,00007	рм	0,00014	рн	0,00071	ро	0,00690
рп	0,00010	рр	0,00001	рс	0,00022	рт	0,00109	ру	0,00270	рф	0,00001	рх	0,00017	рц	0,00005
рч	0,00008	рш	0,00031	рщ	0,00002	ры	0,00143	рь	0,00066	рю	0,00016	ря	0,00100	р	0,00060
са	0,00200	сб	0,00005	св	0,00203	сг	0,00002	сд	0,00027	се	0,00382	сж	0,00002	сз	0,00002
си	0,00135	ск	0,00347	сл	0,00258	см	0,00083	сн	0,00114	со	0,00239	сп	0,00121	ср	0,00011
сс	0,00069	ст	0,01017	су	0,00065	сф	0,00000	сх	0,00020	сц	0,00003	сч	0,00016	сш	0,00007
сы	0,00035	сь	0,00299	сю	0,00020	ся	0,00433	с	0,00288	та	0,00572	тб	0,00001	тв	0,00191
тг	0,00001	тд	0,00011	те	0,00502	тз	0,00001	ти	0,00296	тк	0,00050	тл	0,00018	тм	0,00001
тн	0,00101	то	0,01268	тп	0,00010	тр	0,00253	тс	0,00132	тт	0,00005	ту	0,00173	тф	0,00001
тх	0,00002	тц	0,00009	тч	0,00013	тш	0,00000	тщ	0,00002	ты	0,00155	ть	0,00588	тю	0,00003
тя	0,00054	т	0,00616	уа	0,00001	уб	0,00091	ув	0,00076	уг	0,00119	уд	0,00232	уе	0,00017
уж	0,00166	уз	0,00046	уй	0,00000	уй	0,00008	ук	0,00105	ул	0,00144	ум	0,00131	ун	0,00029
уо	0,00001	уп	0,00073	ур	0,00057	ус	0,00114	ут	0,00147	уф	0,00000	ух	0,00057	уц	0,00001
уч	0,00085	уш	0,00104	ущ	0,00015	уэ	0,00000	ую	0,00103	уя	0,00006	у	0,00668	фа	0,00007
фе	0,00020	фи	0,00011	фл	0,00002	фм	0,00002	фо	0,00008	фр	0,00007	фт	0,00002	фу	0,00003
фф	0,00001	ха	0,00081	хв	0,00030	хг	0,00000	хе	0,00009	хи	0,00022	хл	0,00023	хм	0,00002
хн	0,00017	хо	0,00203	хр	0,00015	хс	0,00003	хт	0,00003	ху	0,00037	хш	0,00001	хъ	0,00001
хэ	0,00000	х	0,00323	ца	0,00081	цв	0,00008	це	0,00101	ци	0,00014	цк	0,00011	цм	0,00002
цо	0,00021	цу	0,00029	цы	0,00036	ць	0,00004	цю	0,00004	ц	0,00077	ча	0,00207	че	0,00318
чи	0,00184	чк	0,00043	чл	0,00001	чм	0,00001	чн	0,00061	чо	0,00007	чр	0,00006	чс	0,00001
чт	0,00345	чу	0,00103	чц	0,00000	чш	0,00016	чь	0,00030	ч	0,00050	ша	0,00091	шв	0,00005
ше	0,00228	ши	0,00236	шк	0,00071	шл	0,00035	шм	0,00004	шн	0,00048	шо	0,00023	шп	0,00006
шт	0,00009	шу	0,00037	шь	0,00054	ш	0,00010	ща	0,00023	шв	0,00000	ше	0,00121	ши	0,00058
щн	0,00001	що	0,00000	щр	0,00001	шу	0,00010	щь	0,00001	щ	0,00002	ьб	0,00014	ьв	0,00099
ьг	0,00016	ьд	0,00009	ье	0,00115	ьж	0,00002	ьз	0,00007	ьи	0,00001	ьй	0,00170	ьк	0,00033
ьл	0,00192	ьм	0,00124	ьн	0,00029	ьп	0,00023	ьр	0,00024	ьс	0,00057	ьт	0,00049	ьу	0,00002
ьх	0,00084	ьц	0,00002	ьч	0,00011	ьш	0,00063	ьщ	0,00004	ь	0,00442	ьб	0,00008	ьг	0,00006
ьд	0,00003	ье	0,00063	ьз	0,00016	ьи	0,00007	ьк	0,00147	ьм	0,00032	ьн	0,00095	ьп	0,00000
ьс	0,00092	ьт	0,00007	ьц	0,00009	ьч	0,00004	ьш	0,00036	ьщ	0,00002	ью	0,00028	ья	0,00070
ь	0,01073	эг	0,00002	эд	0,00002	эз	0,00000	эй	0,00001	эк	0,00006	эл	0,00001	эп	0,00001
эс	0,00001	эт	0,00174	эф	0,00002	эх	0,00003	э	0,00003	юб	0,00031	юг	0,00001	юд	0,00029
юе	0,00001	юж	0,00003	юй	0,00001	юк	0,00007	юл	0,00008	юм	0,00006	юн	0,00004	юр	0,00004
юс	0,00010	ют	0,00047	юх	0,00003	юч	0,00009	юш	0,00003	ющ	0,00022	юю	0,00004	ю	0,00383
яб	0,00002	яв	0,00038	яг	0,00009	яд	0,00067	яе	0,00009	яж	0,00011	яз	0,00024	ян	0,00004
яй	0,00006	як	0,00043	ял	0,00071	ям	0,00040	ян	0,00071	яп	0,00005	яр	0,00017	яс	0,00050
ят	0,00152	ях	0,00016	яц	0,00010	яч	0,00019	яш	0,00001	ящ	0,00014	яю	0,00008	яя	0,00009
я	0,01016	а	0,00154	б	0,00711	в	0,01511	г	0,00463	д	0,00766	е	0,00402	ж	0,00207
з	0,00477	и	0,01025	й	0,00000	к	0,01062	л	0,00286	м	0,00534	н	0,01727	о	0,00959
п	0,01605	р	0,00375	с	0,01617	т	0,00890	у	0,00430	ф	0,00040	х	0,00207	ц	0,00050
ч	0,00615	ш	0,00118	щ	0,00017	э	0,00172	ю	0,00003	я	0,00139				

Табл.2.3 – Частоти біграм з пробілом

аа	0,00018	аб	0,00166	ав	0,00517	аг	0,00130	ад	0,00291	ае	0,00178	аж	0,00141	аз	0,00464
аи	0,00131	ай	0,00087	ак	0,00937	ал	0,00917	ам	0,00476	ан	0,00621	ао	0,00103	ап	0,00289
ар	0,00372	ас	0,00612	ат	0,00661	ау	0,00061	аф	0,00011	ах	0,00137	ац	0,00021	ач	0,00173
аш	0,00119	ащ	0,00032	аз	0,00029	аю	0,00087	ая	0,00258	ба	0,00122	бб	0,00002	бв	0,00008
бг	0,00001	бд	0,00003	бе	0,00249	бж	0,00003	бз	0,00003	би	0,00056	бк	0,00023	бл	0,00096
бм	0,00007	бн	0,00027	бо	0,00254	бп	0,00003	бр	0,00154	бс	0,00018	бт	0,00006	бу	0,00181
бх	0,00003	бц	0,00001	бч	0,00004	бш	0,00002	бщ	0,00010	бы	0,00454	бь	0,00017	бэ	0,00003
бю	0,00000	бя	0,00062	ва	0,00661	вб	0,00025	вв	0,00039	вг	0,00031	вд	0,00082	ве	0,00616
вж	0,00004	вз	0,00059	ви	0,00472	вй	0,00001	вк	0,00095	вл	0,00086	vm	0,00050	вн	0,00163
во	0,00796	вп	0,00084	вр	0,00095	вс	0,00478	вт	0,00054	ву	0,00096	вф	0,00000	вх	0,00011
вц	0,00004	вч	0,00027	вш	0,00175	вщ	0,00007	вы	0,00286	вэ	0,00020	вз	0,00017	вю	0,00000
вя	0,00036	га	0,00117	гб	0,00003	гв	0,00011	гг	0,00001	гд	0,00151	ге	0,00037	гз	0,00008
ги	0,00064	гк	0,00013	гл	0,00176	гм	0,00005	гн	0,00051	го	0,01070	гп	0,00011	гр	0,00097
гс	0,00013	гт	0,00006	гу	0,00086	гч	0,00003	гш	0,00000	гя	0,00001	да	0,00518	дб	0,00013
дв	0,00109	дг	0,00007	дд	0,00007	де	0,00563	дж	0,00004	дз	0,00011	ди	0,00268	дк	0,00050
дл	0,00059	дм	0,00013	дн	0,00293	до	0,00431	дп	0,00019	др	0,00142	дс	0,00044	дт	0,00055

ду	0,00236	дф	0,00000	дх	0,00005	дц	0,00029	дч	0,00009	дш	0,00012	ды	0,00072	дь	0,00098
дэ	0,00001	дю	0,00002	дя	0,00038	еа	0,00021	еб	0,00302	ев	0,00431	ег	0,00470	ед	0,00408
её	0,00206	еж	0,00130	ез	0,00215	ей	0,00134	ей	0,00205	ек	0,00242	ел	0,00712	ем	0,00556
ен	0,00862	ео	0,00124	еп	0,00347	ер	0,00693	ес	0,00765	ет	0,00779	еу	0,00064	еф	0,00004
ех	0,00114	ец	0,00115	еч	0,00222	еш	0,00127	ещ	0,00098	ез	0,00024	ею	0,00051	ея	0,00036
жа	0,00125	жб	0,00005	жв	0,00004	жг	0,00000	жд	0,00081	же	0,00451	жж	0,00002	жз	0,00001
жи	0,00158	жк	0,00018	жл	0,00003	жм	0,00004	жн	0,00127	жо	0,00007	жп	0,00004	жр	0,00002
жс	0,00012	жт	0,00004	жу	0,00029	жц	0,00007	жч	0,00004	жь	0,00005	жэ	0,00002	жя	0,00002
за	0,00660	зб	0,00016	зв	0,00116	зг	0,00039	зд	0,00106	зе	0,00066	зж	0,00015	зз	0,00008
зи	0,00043	зк	0,00025	зл	0,00027	зм	0,00027	зн	0,00217	зо	0,00081	зп	0,00021	зр	0,00026
зс	0,00016	зт	0,00011	зу	0,00039	зх	0,00003	зц	0,00002	зч	0,00010	зш	0,00002	зы	0,00056
зь	0,00028	зэ	0,00002	зю	0,00001	зя	0,00039	иа	0,00024	иб	0,00161	ив	0,00558	иг	0,00118
ид	0,00291	ие	0,00282	иж	0,00053	из	0,00293	ии	0,00148	ий	0,00153	ик	0,00414	ил	0,00601
им	0,00363	ин	0,00597	ио	0,00148	ип	0,00258	ир	0,00145	ис	0,00554	ит	0,00573	иу	0,00061
иф	0,00008	их	0,00212	иц	0,00148	ич	0,00242	иш	0,00065	ищ	0,00019	из	0,00020	ию	0,00029
ия	0,00116	йа	0,00023	йб	0,00039	йв	0,00062	йг	0,00032	йд	0,00065	йе	0,00014	йж	0,00017
йз	0,00023	йи	0,00033	йк	0,00072	йл	0,00018	йм	0,00039	йн	0,00085	йо	0,00044	йп	0,00071
йр	0,00024	йс	0,00116	йт	0,00063	йу	0,00015	йф	0,00004	йх	0,00016	йц	0,00009	йч	0,00039
йш	0,00019	йэ	0,00004	йю	0,00001	йя	0,00007	ка	0,01033	кб	0,00054	кв	0,00085	кг	0,00013
кд	0,00024	ке	0,00075	кж	0,00018	кз	0,00016	ки	0,00464	кк	0,00037	кл	0,00074	км	0,00017
кн	0,00147	ко	0,01138	кп	0,00050	кр	0,00225	кс	0,00069	кт	0,00105	ку	0,00306	кф	0,00001
кх	0,00007	кц	0,00006	кч	0,00031	кш	0,00005	кщ	0,00001	кэ	0,00007	кя	0,00006	ла	0,00652
лб	0,00034	лв	0,00081	лг	0,00052	лд	0,00043	ле	0,00538	лж	0,00030	лз	0,00016	ли	0,00732
лк	0,00108	лл	0,00021	лм	0,00016	лн	0,00101	ло	0,00885	лп	0,00087	лр	0,00014	лс	0,00255
лт	0,00043	лу	0,00185	лф	0,00001	лх	0,00008	лц	0,00001	лч	0,00047	лш	0,00004	лщ	0,00003
лы	0,00094	ль	0,00438	лэ	0,00004	лю	0,00104	ля	0,00210	ма	0,00289	мб	0,00034	мв	0,00083
мг	0,00024	мд	0,00046	ме	0,00435	мж	0,00012	мз	0,00020	ми	0,00415	мй	0,00000	мк	0,00070
мл	0,00050	мм	0,00035	мн	0,00260	мо	0,00444	мп	0,00102	мр	0,00024	мс	0,00104	мт	0,00044
му	0,00279	мф	0,00002	мх	0,00012	мц	0,00006	мч	0,00037	мш	0,00006	мщ	0,00001	мы	0,00093
мь	0,00007	мэ	0,00009	мя	0,00059	на	0,01160	нб	0,00028	нв	0,00042	нг	0,00019	нд	0,00052
не	0,01243	нж	0,00003	нз	0,00017	ни	0,00810	нк	0,00068	нл	0,00007	нм	0,00009	нн	0,00342
но	0,01116	нп	0,00049	нр	0,00017	нс	0,00070	нт	0,00043	ну	0,00341	нф	0,00012	нх	0,00004
нц	0,00031	нч	0,00030	нш	0,00001	нщ	0,00006	ны	0,00386	нь	0,00175	нэ	0,00003	ню	0,00024
ня	0,00155	оа	0,00020	об	0,00508	ов	0,01102	ог	0,00652	од	0,00699	ое	0,00313	ож	0,00283
оз	0,00242	ои	0,00248	ой	0,00332	ок	0,00396	ол	0,00755	ом	0,00656	он	0,00879	оо	0,00176
оп	0,00388	ор	0,00755	ос	0,00909	от	0,00858	оу	0,00092	оф	0,00020	ох	0,00090	оц	0,00013
оч	0,00332	ош	0,00127	ощ	0,00020	оэ	0,00030	ою	0,00158	оя	0,00092	па	0,00205	пв	0,00001
пг	0,00001	пд	0,00000	пе	0,00293	пз	0,00000	пи	0,00136	пк	0,00021	пл	0,00079	пн	0,00013
по	0,01025	пп	0,00003	пр	0,00731	пс	0,00003	пт	0,00007	пу	0,00088	пц	0,00009	пч	0,00007
пш	0,00003	пщ	0,00001	пы	0,00027	пь	0,00011	пя	0,00028	ра	0,00742	рб	0,00016	рв	0,00033
рг	0,00017	рд	0,00038	ре	0,00655	рж	0,00026	рз	0,00008	ри	0,00542	рк	0,00057	рл	0,00011
рм	0,00022	рн	0,00092	ро	0,00841	рп	0,00018	рр	0,00003	рс	0,00036	рт	0,00134	ру	0,00322
рф	0,00001	рх	0,00019	рц	0,00003	рч	0,00010	рш	0,00042	рщ	0,00006	ры	0,00180	рь	0,00080
рэ	0,00001	рю	0,00019	ря	0,00125	са	0,00235	сб	0,00025	св	0,00263	сг	0,00013	сд	0,00047
се	0,00457	сж	0,00008	сз	0,00012	си	0,00175	ск	0,00439	сл	0,00293	см	0,00105	сн	0,00178
со	0,00302	сп	0,00183	ср	0,00020	сс	0,00107	ст	0,01225	су	0,00087	сф	0,00002	сх	0,00030
сц	0,00004	сч	0,00030	сш	0,00016	сщ	0,00000	сы	0,00047	сь	0,00364	сэ	0,00003	сю	0,00021
ся	0,00516	та	0,00696	тб	0,00035	тв	0,00302	тг	0,00023	тд	0,00049	те	0,00603	тж	0,00017
тз	0,00021	ти	0,00406	тк	0,00110	тл	0,00041	тм	0,00023	тн	0,00208	то	0,01579	тп	0,00077
тр	0,00345	тс	0,00220	тт	0,00044	ту	0,00225	тф	0,00001	тх	0,00007	тц	0,00014	тч	0,00043
тш	0,00002	тщ	0,00003	ты	0,00186	ть	0,00708	тэ	0,00011	тю	0,00004	тя	0,00062	уа	0,00012
уб	0,00139	ув	0,00165	уг	0,00157	уд	0,00319	уе	0,00035	уж	0,00193	уз	0,00084	уй	0,00064
уй	0,00008	ук	0,00176	ул	0,00189	ум	0,00188	ун	0,00148	уо	0,00036	уп	0,00173	ур	0,00083
ус	0,00216	ут	0,00217	уу	0,00014	уф	0,00004	ух	0,00073	уц	0,00003	уч	0,00148	уш	0,00118
ущ	0,00019	уэ	0,00007	ую	0,00117	уя	0,00010	фа	0,00010	фб	0,00000	фв	0,00000	фе	0,00024
фи	0,00014	фл	0,00003	фм	0,00002	фо	0,00010	фр	0,00007	фс	0,00000	фт	0,00003	фу	0,00003
фф	0,00001	фч	0,00000	фы	0,00000	ха	0,00106	хб	0,00013	хв	0,00075	хг	0,00013	хд	0,00015
хе	0,00015	хж	0,00004	хз	0,00010	хи	0,00057	хк	0,00034	хл	0,00037	хм	0,00018	хн	0,00048
хо	0,00260	хп	0,00041	хр	0,00032	хс	0,00054	хт	0,00018	ху	0,00059	хф	0,00001	хх	0,00006
хц	0,00003	хч	0,00010	хш	0,00005	хщ	0,00001	хь	0,00001	хэ	0,00006	хя	0,00002	ца	0,00108
цб	0,00003	цв	0,00023	цг	0,00001	цд	0,00003	це	0,00120	цж	0,00001	цз	0,00003	ци	0,00022
цк	0,00021	цл	0,00001	цм	0,00009	цн	0,00007	цо	0,00034	цп	0,00010	цр	0,00004	цс	0,00010
цт	0,00005	цу	0,00024	цф	0,00000	цх	0,00001	цч	0,00001	цш	0,00000	цы	0,00040	ць	0,00005
цэ	0,00001	цю	0,00005	ця	0,00003	ча	0,00245	чб	0,00002	чв	0,00006	чг	0,00000	чд	0,00001
че	0,00406	чж	0,00000	чз	0,00003	чи	0,00222	чк	0,00055	чл	0,00002	чм	0,00001	чн	0,00073
чо	0,00008	чп	0,00008	чр	0,00009	чс	0,00007	чт	0,00408	чу	0,00124	чф	0,00000	чх	0,00001
чц	0,00000	чч	0,00001	чш	0,00024	чь	0,00032	чэ	0,00001	чя	0,00001	ша	0,00128	шв	0,00006
шд	0,00001	ше	0,00265	шз	0,00000	ши	0,00295	шк	0,00091	шл	0,00048	шм	0,00004	шн	0,00062

шо	0,00030	шп	0,00009	шр	0,00000	шс	0,00003	шт	0,00008	шу	0,00045	шь	0,00066	ща	0,00028
шб	0,00000	шв	0,00001	шд	0,00000	ше	0,00149	ши	0,00062	шк	0,00000	шн	0,00002	шр	0,00001
шт	0,00000	шу	0,00013	шь	0,00001	ша	0,00004	ыб	0,00041	ыв	0,00179	ыг	0,00032	ыд	0,00032
ые	0,00168	ыж	0,00005	ыз	0,00021	ыи	0,00052	ый	0,00209	ык	0,00068	ыл	0,00234	ым	0,00173
ын	0,00095	ьо	0,00029	ып	0,00078	ыр	0,00039	ыс	0,00125	ыт	0,00082	ыу	0,00016	ыф	0,00000
ых	0,00109	ыц	0,00003	ыч	0,00027	ыш	0,00067	ыщ	0,00005	ыз	0,00005	ыя	0,00008	ба	0,00012
ьб	0,00053	ьв	0,00130	ьг	0,00030	ьд	0,00066	ье	0,00129	ьж	0,00008	ьз	0,00053	би	0,00105
ьк	0,00248	ьл	0,00021	ьм	0,00076	ьн	0,00267	ьо	0,00075	ьп	0,00123	ьр	0,00024	ьс	0,00223
ьт	0,00078	ьу	0,00028	ьф	0,00002	ьх	0,00013	ьц	0,00014	ьч	0,00065	ьш	0,00046	ьщ	0,00003
ьэ	0,00012	ьо	0,00028	ья	0,00087	эг	0,00002	эд	0,00001	эй	0,00001	эк	0,00007	эл	0,00002
эн	0,00000	эп	0,00000	эс	0,00001	эт	0,00202	эф	0,00001	эх	0,00004	эч	0,00000	ээ	0,00000
юа	0,00008	юб	0,00043	юв	0,00037	юг	0,00017	юд	0,00054	юе	0,00009	юж	0,00012	юз	0,00010
юи	0,00041	юк	0,00051	юл	0,00020	юм	0,00025	юн	0,00047	юо	0,00022	юп	0,00044	юр	0,00017
юс	0,00060	ют	0,00087	юу	0,00010	юф	0,00003	юх	0,00013	юц	0,00002	юч	0,00028	юш	0,00011
ющ	0,00028	юз	0,00004	юю	0,00004	юя	0,00003	яа	0,00014	яб	0,00052	яв	0,00168	яг	0,00048
яд	0,00134	яе	0,00044	яж	0,00031	яз	0,00063	яи	0,00082	яй	0,00009	як	0,00138	ял	0,00101
ям	0,00075	ян	0,00250	яо	0,00070	яп	0,00114	яр	0,00039	яс	0,00180	ят	0,00247	яу	0,00023
яф	0,00002	ях	0,00033	яц	0,00013	яч	0,00059	яш	0,00008	ящ	0,00018	яэ	0,00011	яю	0,00008
яя	0,00025														

Табл.2.4 – Частоти біграм без пробілу

аб	0,00079	ав	0,00274	аг	0,00057	ад	0,00165	ае	0,00108	аж	0,00108	аз	0,00346	аи	0,00007
ай	0,00073	ак	0,00653	ал	0,00749	ам	0,00344	ан	0,00347	ао	0,00002	ап	0,00089	ар	0,00270
ас	0,00342	ат	0,00457	ау	0,00008	аф	0,00006	ах	0,00094	ац	0,00014	ач	0,00091	аш	0,00088
ац	0,00023	аэ	0,00001	аю	0,00072	ая	0,00196	а	0,01626	ба	0,00105	бб	0,00001	бв	0,00003
бг	0,00000	бд	0,00002	бе	0,00210	бж	0,00002	бз	0,00000	би	0,00051	бк	0,00018	бл	0,00085
бм	0,00005	бн	0,00018	бо	0,00207	бп	0,00000	бр	0,00130	бс	0,00012	бт	0,00001	бу	0,00147
бх	0,00003	бц	0,00001	бч	0,00002	бш	0,00001	бщ	0,00008	бы	0,00378	бь	0,00014	бю	0,00000
бя	0,00050	б	0,00039	ва	0,00572	вб	0,00001	вв	0,00008	вг	0,00001	вд	0,00032	ве	0,00504
вж	0,00000	вз	0,00039	ви	0,00357	вк	0,00026	вл	0,00061	vm	0,00024	вн	0,00092	во	0,00644
вп	0,00021	вр	0,00054	вс	0,00325	вт	0,00008	ву	0,00067	вх	0,00002	вц	0,00002	вч	0,00008
вш	0,00145	вщ	0,00004	вы	0,00245	вь	0,00016	во	0,00000	вя	0,00028	в	0,00552	га	0,00092
гб	0,00000	гв	0,00002	гд	0,00119	ге	0,00026	ги	0,00057	гк	0,00009	гл	0,00143	гм	0,00002
гн	0,00032	го	0,00887	гр	0,00089	гс	0,00002	гт	0,00003	гу	0,00070	гч	0,00001	гш	0,00001
г	0,00077	да	0,00436	дб	0,00005	дв	0,00085	дг	0,00001	дд	0,00001	де	0,00452	дж	0,00002
дз	0,00002	ди	0,00222	дк	0,00033	дл	0,00045	дм	0,00007	дн	0,00211	до	0,00353	дп	0,00006
др	0,00108	дс	0,00024	дт	0,00041	ду	0,00195	дх	0,00004	дц	0,00025	дч	0,00002	дш	0,00009
ды	0,00064	дь	0,00081	дю	0,00002	дя	0,00035	д	0,00138	еа	0,00003	еб	0,00147	ев	0,00198
ег	0,00352	ед	0,00266	ее	0,00141	еж	0,00090	ез	0,00120	еи	0,00013	ей	0,00165	ек	0,00112
ел	0,00561	ем	0,00395	ен	0,00559	ео	0,00019	еп	0,00101	ер	0,00549	ес	0,00446	ет	0,00550
еу	0,00006	еф	0,00000	ех	0,00065	ец	0,00093	еч	0,00111	еш	0,00096	ещ	0,00078	ею	0,00044
ея	0,00020	е	0,01827	жа	0,00109	жб	0,00002	жг	0,00000	жд	0,00059	же	0,00362	жж	0,00002
жи	0,00138	жк	0,00016	жл	0,00001	жм	0,00001	жн	0,00090	жо	0,00004	жр	0,00001	жс	0,00006
жу	0,00021	жц	0,00006	жч	0,00002	жъ	0,00004	ж	0,00042	за	0,00547	зб	0,00011	зв	0,00084
зг	0,00025	зд	0,00076	зе	0,00049	зж	0,00010	зз	0,00000	зи	0,00035	зк	0,00011	зл	0,00021
зм	0,00016	зн	0,00173	зо	0,00057	зр	0,00014	зс	0,00001	зт	0,00002	зу	0,00029	зч	0,00004
зш	0,00002	зы	0,00048	зь	0,00023	зю	0,00000	зя	0,00032	з	0,00142	иа	0,00006	иб	0,00052
ив	0,00260	иг	0,00047	ид	0,00163	ие	0,00179	иж	0,00031	из	0,00191	ии	0,00029	ий	0,00120
ик	0,00222	ил	0,00465	им	0,00240	ин	0,00319	ио	0,00006	ип	0,00015	ир	0,00074	ис	0,00270
ит	0,00363	иу	0,00002	иф	0,00003	их	0,00157	иц	0,00117	ич	0,00153	иш	0,00044	ищ	0,00015
ию	0,00023	ия	0,00081	и	0,01819	йв	0,00000	йд	0,00018	йе	0,00000	йз	0,00000	йк	0,00011
йл	0,00001	йм	0,00003	йн	0,00028	йо	0,00006	йр	0,00000	йс	0,00028	йт	0,00020	йц	0,00001
йч	0,00004	йш	0,00009	йя	0,00000	й	0,00688	ка	0,00877	кв	0,00024	кд	0,00000	ке	0,00054
кж	0,00003	кз	0,00000	ки	0,00365	кк	0,00001	кл	0,00057	км	0,00000	кн	0,00058	ко	0,00919
кр	0,00176	кс	0,00009	кт	0,00050	ку	0,00237	кц	0,00002	кш	0,00000	к	0,00544	ла	0,00527
лб	0,00005	лв	0,00005	лг	0,00019	лд	0,00010	ле	0,00399	лж	0,00024	лз	0,00001	ли	0,00584
лк	0,00033	лл	0,00014	лм	0,00001	лн	0,00023	ло	0,00641	лп	0,00010	лс	0,00161	лт	0,00007
лу	0,00135	лф	0,00000	лх	0,00000	лч	0,00010	лш	0,00000	лщ	0,00001	лы	0,00072	ль	0,00352
лю	0,00090	ля	0,00174	л	0,00660	ма	0,00238	мб	0,00001	мг	0,00003	ме	0,00353	ми	0,00292
мк	0,00012	мл	0,00030	мм	0,00002	мн	0,00140	мо	0,00315	мп	0,00004	мр	0,00004	мс	0,00008
мт	0,00001	му	0,00213	мф	0,00000	мц	0,00003	мч	0,00003	мш	0,00000	мщ	0,00000	мы	0,00077
мь	0,00004	мя	0,00040	м	0,00759	на	0,00957	нб	0,00002	нв	0,00000	нг	0,00003	нд	0,00021
не	0,01037	нж	0,00001	нз	0,00002	ни	0,00638	нк	0,00036	нм	0,00000	нн	0,00245	но	0,00918
нр	0,00003	нс	0,00027	нт	0,00024	ну	0,00272	нф	0,00001	нх	0,00000	нц	0,00023	нч	0,00012
нш	0,00000	нщ	0,00004	ны	0,00320	нь	0,00145	ню	0,00017	ня	0,00128	н	0,00388	об	0,00314
ов	0,00695	ог	0,00485	од	0,00476	ое	0,00203	ож	0,00187	оз	0,00134	ои	0,00098	ой	0,00268
ок	0,00213	ол	0,00619	ом	0,00480	он	0,00500	оо	0,00013	оп	0,00123	ор	0,00574	ос	0,00540

от	0,00598	оу	0,00006	оф	0,00013	ох	0,00045	оц	0,00006	оч	0,00184	ош	0,00090	ощ	0,00015
оз	0,00002	ою	0,00137	оя	0,00058	о	0,02220	па	0,00167	пе	0,00241	пи	0,00106	пк	0,00019
пл	0,00066	пн	0,00010	по	0,00859	пп	0,00001	пр	0,00607	пс	0,00003	пт	0,00006	пу	0,00069
пц	0,00007	пч	0,00005	пш	0,00002	пщ	0,00001	пы	0,00020	пь	0,00009	пя	0,00027	п	0,00011
ра	0,00607	рб	0,00008	рв	0,00021	рг	0,00014	рд	0,00029	ре	0,00536	рж	0,00021	рз	0,00004
ри	0,00472	рк	0,00039	рл	0,00007	рм	0,00015	рн	0,00074	ро	0,00707	рп	0,00010	рр	0,00001
рс	0,00023	рт	0,00110	ру	0,00279	рф	0,00000	рх	0,00017	рц	0,00003	рч	0,00008	рш	0,00032
рщ	0,00003	ры	0,00146	рь	0,00067	рю	0,00014	ря	0,00103	р	0,00063	са	0,00196	сб	0,00006
св	0,00205	сг	0,00002	сд	0,00027	се	0,00373	сж	0,00002	сз	0,00002	си	0,00136	ск	0,00346
сл	0,00251	см	0,00083	сн	0,00111	со	0,00232	сп	0,00123	ср	0,00013	сс	0,00069	ст	0,01008
су	0,00065	сф	0,00000	сх	0,00020	сц	0,00002	сч	0,00018	сш	0,00010	сы	0,00039	сь	0,00300
сю	0,00019	ся	0,00432	с	0,00284	та	0,00575	тб	0,00002	тв	0,00194	тг	0,00001	тд	0,00012
те	0,00493	тз	0,00002	ти	0,00308	тк	0,00048	тл	0,00020	тм	0,00001	тн	0,00104	то	0,01269
тп	0,00010	тр	0,00259	тс	0,00130	тт	0,00006	ту	0,00172	тф	0,00000	тх	0,00001	тц	0,00009
тч	0,00013	тш	0,00000	тщ	0,00001	ты	0,00153	ть	0,00598	тю	0,00004	тя	0,00055	т	0,00623
уа	0,00002	уб	0,00088	ув	0,00083	уг	0,00114	уд	0,00239	уе	0,00016	уж	0,00158	уз	0,00049
уй	0,00000	уй	0,00008	ук	0,00103	ул	0,00145	ум	0,00131	ун	0,00031	уо	0,00000	уп	0,00074
ур	0,00059	ус	0,00119	ут	0,00145	уф	0,00000	ух	0,00057	уц	0,00001	уч	0,00090	уш	0,00097
ущ	0,00015	уз	0,00000	ую	0,00102	уя	0,00006	у	0,00685	фа	0,00008	фе	0,00020	фи	0,00013
фл	0,00002	фм	0,00002	фо	0,00009	фр	0,00007	фс	0,00000	фт	0,00002	фу	0,00003	фф	0,00001
фы	0,00000	фь	0,00000	ф	0,00000	ха	0,00080	хв	0,00030	хг	0,00000	хе	0,00008	хи	0,00024
хл	0,00022	хм	0,00004	хн	0,00018	хо	0,00208	хр	0,00014	хс	0,00003	хт	0,00004	ху	0,00037
хш	0,00001	хь	0,00000	хэ	0,00001	х	0,00327	ца	0,00085	цв	0,00009	це	0,00097	ци	0,00012
цк	0,00012	цм	0,00002	цз	0,00023	цу	0,00023	цы	0,00033	ць	0,00004	цю	0,00004	ця	0,00000
ц	0,00082	ча	0,00206	чв	0,00000	че	0,00323	чи	0,00182	чк	0,00045	чл	0,00001	чм	0,00001
чн	0,00062	чо	0,00007	чр	0,00007	чс	0,00001	чт	0,00340	чу	0,00097	чц	0,00000	чш	0,00018
чь	0,00030	ч	0,00050	ша	0,00103	шв	0,00005	ше	0,00222	ши	0,00236	шк	0,00070	шл	0,00038
шм	0,00003	шн	0,00049	шо	0,00025	шп	0,00006	шт	0,00008	шу	0,00037	шь	0,00052	ш	0,00010
ща	0,00028	шв	0,00000	ше	0,00124	щи	0,00057	щн	0,00001	що	0,00000	щр	0,00001	щу	0,00010
щъ	0,00001	щ	0,00002	ьб	0,00016	ьв	0,00106	ьг	0,00015	ьд	0,00011	ье	0,00123	ьж	0,00001
ьз	0,00006	ьи	0,00001	ьй	0,00174	ьк	0,00033	ьл	0,00198	ьм	0,00123	ьн	0,00031	ьп	0,00021
ьр	0,00025	ьс	0,00056	ьт	0,00048	ьу	0,00001	ьх	0,00086	ьц	0,00001	ьч	0,00010	ьш	0,00061
ьщ	0,00004	ья	0,00000	ы	0,00444	ьб	0,00009	ьв	0,00000	ьг	0,00005	ьд	0,00003	ье	0,00059
ьз	0,00015	ьи	0,00007	ьк	0,00151	ьм	0,00031	ьн	0,00106	ьо	0,00000	ьп	0,00000	ьс	0,00090
ьт	0,00007	ьф	0,00000	ьц	0,00011	ьч	0,00003	ьш	0,00035	ьщ	0,00002	ьз	0,00000	ью	0,00030
ья	0,00065	ь	0,01071	эг	0,00001	эд	0,00001	эз	0,00000	эй	0,00002	эк	0,00005	эл	0,00001
эн	0,00000	эп	0,00000	эс	0,00001	эт	0,00166	эф	0,00001	эх	0,00003	э	0,00003	юб	0,00027
юв	0,00000	юг	0,00001	юд	0,00028	юе	0,00001	юж	0,00003	юз	0,00000	юй	0,00000	юк	0,00008
юл	0,00008	юм	0,00004	юн	0,00004	юр	0,00004	юс	0,00011	ют	0,00047	юх	0,00004	юц	0,00000
юч	0,00009	юш	0,00003	ющ	0,00022	юю	0,00002	ю	0,00384	яа	0,00000	яб	0,00002	яв	0,00036
яг	0,00007	яд	0,00064	яе	0,00010	яж	0,00012	яз	0,00020	яи	0,00004	яй	0,00007	як	0,00042
ял	0,00075	ям	0,00038	ян	0,00073	яп	0,00005	яр	0,00015	яс	0,00049	ят	0,00148	ях	0,00017
яц	0,00010	яч	0,00016	яш	0,00001	ящ	0,00012	яю	0,00008	яя	0,00010	я	0,01005	а	0,00161
б	0,00715	в	0,01517	г	0,00459	д	0,00773	е	0,00407	ж	0,00209	з	0,00476	и	0,01021
й	0,00001	к	0,01050	л	0,00273	м	0,00545	н	0,01716	о	0,00963	п	0,01612	р	0,00376
с	0,01618	т	0,00891	у	0,00414	ф	0,00040	х	0,00205	ц	0,00048	ч	0,00611	ш	0,00118
щ	0,00017	э	0,00181	ю	0,00003	я	0,00146								

Табл.2.5 – Частоти біграм з перетином з пробілом

аа	0,00016	аб	0,00171	ав	0,00506	аг	0,00120	ад	0,00293	ае	0,00180	аж	0,00148	аз	0,00467
аи	0,00133	ай	0,00087	ак	0,00921	ал	0,00934	ам	0,00477	ан	0,00634	ао	0,00107	ап	0,00285
ар	0,00364	ас	0,00602	ат	0,00653	ау	0,00058	аф	0,00014	ах	0,00130	ац	0,00021	ач	0,00173
аш	0,00122	ащ	0,00029	аз	0,00027	аю	0,00087	ая	0,00257	ба	0,00127	бб	0,00002	бв	0,00007
бг	0,00000	бд	0,00003	бе	0,00255	бж	0,00003	бз	0,00002	би	0,00065	бк	0,00023	бл	0,00102
бм	0,00008	бн	0,00025	бо	0,00253	бп	0,00003	бр	0,00157	бс	0,00016	бт	0,00005	бу	0,00178
бх	0,00004	бц	0,00001	бч	0,00004	бш	0,00002	бщ	0,00009	бы	0,00453	бь	0,00017	бэ	0,00006
бю	0,00000	бя	0,00061	ва	0,00690	вб	0,00024	вв	0,00039	вг	0,00031	вд	0,00077	ве	0,00616
вж	0,00005	вз	0,00064	ви	0,00455	вій	0,00000	вк	0,00091	вл	0,00088	вм	0,00048	вн	0,00171
во	0,00809	вп	0,00085	вр	0,00090	вс	0,00469	вт	0,00055	ву	0,00093	вф	0,00001	вх	0,00011
вц	0,00005	вч	0,00028	вш	0,00185	вщ	0,00006	вы	0,00293	вь	0,00019	вэ	0,00018	вю	0,00000
вя	0,00036	га	0,00111	гб	0,00004	гв	0,00012	гг	0,00002	гд	0,00148	ге	0,00033	гз	0,00007
ги	0,00073	гк	0,00014	гл	0,00172	гм	0,00005	гн	0,00048	го	0,01071	гп	0,00012	гр	0,00109
гс	0,00014	гт	0,00006	гу	0,00088	гх	0,00000	гч	0,00003	гш	0,00001	гэ	0,00000	гя	0,00000
да	0,00524	дб	0,00012	дв	0,00110	дг	0,00006	дд	0,00007	де	0,00545	дж	0,00004	дз	0,00011
ди	0,00281	дк	0,00051	дл	0,00058	дм	0,00012	дн	0,00285	до	0,00432	дп	0,00019	др	0,00135
дс	0,00045	дт	0,00055	ду	0,00237	дф	0,00000	дх	0,00006	дц	0,00031	дч	0,00007	дш	0,00012
ды	0,00077	дь	0,00097	дэ	0,00002	дю	0,00003	дя	0,00044	еа	0,00023	еб	0,00291	ев	0,00431

ег	0,00480	ед	0,00424	ее	0,00215	еж	0,00134	ез	0,00217	ей	0,00142	ей	0,00198	ек	0,00242
ел	0,00724	ем	0,00570	ен	0,00863	ео	0,00137	еп	0,00352	ер	0,00705	ес	0,00760	ет	0,00779
еу	0,00069	еф	0,00004	ех	0,00116	ец	0,00116	еч	0,00216	еш	0,00126	ещ	0,00095	ез	0,00022
ею	0,00053	ея	0,00043	жа	0,00130	жб	0,00004	жв	0,00004	жг	0,00001	жд	0,00076	же	0,00435
жж	0,00002	жз	0,00001	жи	0,00168	жк	0,00021	жл	0,00003	жм	0,00004	жн	0,00118	жо	0,00008
жп	0,00003	жр	0,00002	жс	0,00010	жт	0,00004	жу	0,00026	жх	0,00001	жц	0,00007	жч	0,00004
жщ	0,00000	жь	0,00005	жэ	0,00002	жя	0,00001	за	0,00657	зб	0,00017	зв	0,00116	зг	0,00037
зд	0,00102	зе	0,00063	зж	0,00013	зз	0,00009	зи	0,00045	зк	0,00025	зл	0,00027	зм	0,00026
зн	0,00226	зо	0,00077	зп	0,00018	зр	0,00024	зс	0,00016	зт	0,00014	зу	0,00039	зф	0,00000
зх	0,00003	зц	0,00002	зч	0,00009	зш	0,00003	зы	0,00058	зь	0,00027	зэ	0,00002	зю	0,00000
зя	0,00039	иа	0,00025	иб	0,00152	ив	0,00538	иг	0,00117	ид	0,00295	ие	0,00271	иж	0,00059
из	0,00299	ии	0,00152	ий	0,00144	ик	0,00404	ил	0,00588	им	0,00352	ин	0,00593	ио	0,00146
ип	0,00264	ир	0,00141	ис	0,00542	ит	0,00540	иу	0,00060	иф	0,00007	их	0,00215	иц	0,00147
ич	0,00250	иш	0,00068	ищ	0,00022	из	0,00021	ию	0,00027	ия	0,00113	йа	0,00023	йб	0,00038
йв	0,00060	йг	0,00035	йд	0,00067	йе	0,00012	йж	0,00017	йз	0,00022	йи	0,00041	йк	0,00071
йл	0,00016	йм	0,00041	йн	0,00092	йо	0,00043	йп	0,00078	йр	0,00029	йс	0,00125	йт	0,00061
йу	0,00015	йф	0,00003	йх	0,00014	йц	0,00007	йч	0,00038	йш	0,00020	йщ	0,00000	йэ	0,00004
йю	0,00001	йя	0,00007	ка	0,01056	кб	0,00053	кв	0,00088	кг	0,00016	кд	0,00024	ке	0,00075
кж	0,00018	кз	0,00016	ки	0,00482	кк	0,00036	кл	0,00075	км	0,00019	кн	0,00148	ко	0,01133
кп	0,00050	кр	0,00225	кс	0,00074	кт	0,00098	ку	0,00300	кф	0,00001	кх	0,00008	кц	0,00006
кч	0,00031	кш	0,00004	кщ	0,00000	кэ	0,00008	кя	0,00005	ла	0,00636	лб	0,00033	лв	0,00084
лг	0,00046	лд	0,00042	ле	0,00519	лж	0,00031	лз	0,00018	ли	0,00749	лк	0,00100	лл	0,00027
лм	0,00015	лн	0,00094	ло	0,00864	лп	0,00089	лр	0,00013	лс	0,00259	лт	0,00043	лу	0,00181
лф	0,00001	лх	0,00008	лц	0,00001	лч	0,00049	лш	0,00005	лщ	0,00002	лы	0,00087	ль	0,00422
лэ	0,00005	лю	0,00108	ля	0,00213	ма	0,00294	мб	0,00034	мв	0,00080	мг	0,00029	мд	0,00048
ме	0,00439	мж	0,00012	мз	0,00023	ми	0,00419	мй	0,00000	мк	0,00070	мл	0,00048	мм	0,00032
мн	0,00259	мо	0,00436	мп	0,00101	мр	0,00024	мс	0,00103	мт	0,00043	му	0,00277	мф	0,00002
мх	0,00011	мц	0,00007	мч	0,00039	мш	0,00008	мщ	0,00001	мы	0,00092	мь	0,00005	мэ	0,00007
мю	0,00000	мя	0,00056	на	0,01151	нб	0,00028	нв	0,00044	нг	0,00018	нд	0,00053	не	0,01250
нж	0,00004	нз	0,00016	ни	0,00796	нк	0,00064	нл	0,00005	нм	0,00009	нн	0,00342	но	0,01129
нп	0,00053	нр	0,00016	нс	0,00074	нт	0,00048	ну	0,00342	нф	0,00013	нх	0,00005	нц	0,00027
нч	0,00025	нш	0,00003	нщ	0,00006	ны	0,00384	нь	0,00174	нэ	0,00003	ню	0,00021	ня	0,00161
оа	0,00019	об	0,00508	ов	0,01104	ог	0,00645	од	0,00686	ое	0,00307	ож	0,00282	оз	0,00242
ои	0,00244	ой	0,00322	ок	0,00385	ол	0,00779	ом	0,00657	он	0,00884	оо	0,00171	оп	0,00393
ор	0,00753	ос	0,00905	от	0,00872	оу	0,00090	оф	0,00018	ох	0,00090	оц	0,00014	оч	0,00334
ош	0,00126	ощ	0,00019	оэ	0,00032	ою	0,00164	оя	0,00095	па	0,00201	пб	0,00000	пв	0,00001
пг	0,00001	пд	0,00000	пе	0,00289	пз	0,00000	пи	0,00129	пк	0,00023	пл	0,00080	пм	0,00000
пн	0,00014	по	0,01030	пп	0,00003	пр	0,00728	пс	0,00004	пт	0,00008	пу	0,00082	пц	0,00008
пч	0,00006	пш	0,00003	пщ	0,00001	пы	0,00023	пь	0,00011	пя	0,00032	ра	0,00729	рб	0,00014
рв	0,00031	рг	0,00019	рд	0,00037	ре	0,00644	рж	0,00025	рз	0,00006	ри	0,00571	рк	0,00056
рл	0,00010	рм	0,00019	рн	0,00097	ро	0,00852	рп	0,00021	рр	0,00002	рс	0,00035	рт	0,00134
ру	0,00337	рф	0,00001	рх	0,00020	рц	0,00004	рч	0,00012	рш	0,00040	рщ	0,00004	ры	0,00175
рь	0,00080	рэ	0,00001	рю	0,00018	ря	0,00124	са	0,00237	сб	0,00023	св	0,00266	сг	0,00013
сд	0,00048	се	0,00452	сж	0,00006	сз	0,00014	си	0,00173	ск	0,00442	сл	0,00309	см	0,00112
сн	0,00180	со	0,00295	сп	0,00179	ср	0,00022	сс	0,00115	ст	0,01236	су	0,00088	сф	0,00001
сх	0,00027	сц	0,00003	сч	0,00031	сш	0,00015	сщ	0,00000	сы	0,00047	сь	0,00359	сэ	0,00003
сю	0,00023	ся	0,00520	та	0,00696	тб	0,00035	тв	0,00300	тг	0,00024	тд	0,00048	те	0,00610
тж	0,00015	тз	0,00018	ти	0,00416	тк	0,00112	тл	0,00040	тм	0,00023	тн	0,00202	то	0,01567
тп	0,00079	тр	0,00329	тс	0,00216	тт	0,00053	ту	0,00222	тф	0,00001	тх	0,00008	тц	0,00014
тч	0,00044	тш	0,00003	тщ	0,00002	ты	0,00184	ть	0,00716	тэ	0,00011	тю	0,00004	тя	0,00073
уа	0,00013	уб	0,00133	ув	0,00169	уг	0,00154	уд	0,00320	уе	0,00035	уж	0,00201	уз	0,00080
уй	0,00070	уй	0,00010	ук	0,00178	ул	0,00184	ум	0,00190	ун	0,00142	уо	0,00035	уп	0,00171
ур	0,00082	ус	0,00216	ут	0,00223	уу	0,00016	уф	0,00003	ух	0,00076	уц	0,00003	уч	0,00151
уш	0,00121	ущ	0,00019	уэ	0,00007	ую	0,00122	уя	0,00014	фа	0,00009	фб	0,00000	фв	0,00000
фе	0,00023	фи	0,00016	фл	0,00002	фм	0,00003	фо	0,00011	фр	0,00008	фс	0,00000	фт	0,00003
фу	0,00004	фф	0,00001	фч	0,00000	фы	0,00000	фь	0,00000	ха	0,00099	хб	0,00014	хв	0,00068
хг	0,00012	хд	0,00018	хе	0,00016	хж	0,00004	хз	0,00011	хи	0,00062	хк	0,00030	хл	0,00039
хм	0,00017	хн	0,00052	хо	0,00269	хп	0,00041	хр	0,00028	хс	0,00049	хт	0,00020	ху	0,00054
хф	0,00001	хх	0,00005	хц	0,00003	хч	0,00011	хш	0,00005	хщ	0,00001	хь	0,00001	хэ	0,00005
хя	0,00003	ца	0,00105	цб	0,00003	цв	0,00020	цг	0,00001	цд	0,00003	це	0,00118	цж	0,00001
цз	0,00002	ци	0,00023	цк	0,00020	цл	0,00001	цм	0,00008	цн	0,00009	цо	0,00035	цп	0,00010
цр	0,00003	цс	0,00010	цт	0,00005	цу	0,00031	цф	0,00000	цх	0,00001	цц	0,00000	цч	0,00002
цш	0,00001	цы	0,00039	ць	0,00005	цэ	0,00001	цю	0,00005	ця	0,00002	ча	0,00247	чб	0,00002
чв	0,00005	чг	0,00001	чд	0,00003	че	0,00387	чж	0,00001	чз	0,00002	чи	0,00223	чк	0,00057
чл	0,00002	чм	0,00002	чн	0,00080	чо	0,00011	чп	0,00008	чр	0,00010	чс	0,00007	чт	0,00410
чу	0,00119	чф	0,00000	чх	0,00001	чц	0,00000	чч	0,00001	чш	0,00023	чь	0,00036	чэ	0,00001
чя	0,00001	ша	0,00124	шб	0,00000	шв	0,00006	шг	0,00000	шд	0,00001	ше	0,00266	шж	0,00000
шз	0,00000	ши	0,00284	шк	0,00085	шл	0,00046	шм	0,00004	шн	0,00059	шо	0,00030	шп	0,00009

шр	0,00000	шс	0,00002	шт	0,00010	шу	0,00044	шь	0,00063	шэ	0,00000	шы	0,00000	ша	0,00033
шб	0,00000	шв	0,00000	шд	0,00000	ше	0,00149	ши	0,00069	шк	0,00000	шн	0,00002	шо	0,00000
шп	0,00000	шр	0,00001	шт	0,00000	шу	0,00012	шь	0,00001	ша	0,00003	ыб	0,00042	ыв	0,00177
ыг	0,00032	ыд	0,00033	ые	0,00164	ыж	0,00004	ыз	0,00023	ыи	0,00051	ый	0,00208	ык	0,00071
ыл	0,00244	ым	0,00169	ын	0,00095	ыо	0,00025	ып	0,00079	ыр	0,00039	ыс	0,00120	ыт	0,00086
ыу	0,00016	ыф	0,00000	ых	0,00111	ыц	0,00002	ыч	0,00026	ыш	0,00075	ыщ	0,00005	ыз	0,00005
ью	0,00000	ья	0,00006	ба	0,00013	бб	0,00050	бв	0,00131	бг	0,00032	бд	0,00065	бе	0,00127
бж	0,00007	бз	0,00050	би	0,00105	бк	0,00265	бл	0,00022	бм	0,00077	бн	0,00265	бо	0,00077
бп	0,00118	бр	0,00023	бс	0,00220	бт	0,00079	бу	0,00027	бф	0,00001	бх	0,00014	бц	0,00016
бч	0,00061	бш	0,00047	бщ	0,00004	бэ	0,00014	бю	0,00036	бя	0,00091	эб	0,00000	эв	0,00000
эг	0,00002	эд	0,00001	эз	0,00000	эй	0,00002	эк	0,00006	эл	0,00001	эн	0,00001	эо	0,00000
эп	0,00001	эс	0,00001	эт	0,00199	эф	0,00001	эх	0,00004	эч	0,00000	эш	0,00000	юа	0,00008
юб	0,00047	юв	0,00034	юг	0,00014	юд	0,00052	юе	0,00007	юж	0,00012	юз	0,00010	юи	0,00037
юй	0,00000	юк	0,00050	юл	0,00019	юм	0,00023	юн	0,00045	юо	0,00021	юп	0,00042	юр	0,00017
юс	0,00056	ют	0,00086	юу	0,00009	юф	0,00002	юх	0,00013	юц	0,00003	юч	0,00029	юш	0,00010
ющ	0,00027	юз	0,00004	юю	0,00003	юя	0,00005	яа	0,00015	яб	0,00055	яв	0,00172	яг	0,00042
яд	0,00132	яе	0,00039	яж	0,00030	яз	0,00061	яи	0,00080	яй	0,00008	як	0,00137	ял	0,00104
ям	0,00075	ян	0,00239	яо	0,00068	яп	0,00115	яр	0,00042	яс	0,00177	ят	0,00238	яу	0,00022
яф	0,00001	ях	0,00032	яц	0,00014	яч	0,00057	яш	0,00010	ящ	0,00015	яэ	0,00009	яю	0,00011
яя	0,00022														

Табл.2.6 – Частоти біграм з перетином без пробілу

3. Значення ентропії

H_1 – ентропія букв алфавіту, H_2 – ентропія біграм, H_2^* – ентропія біграм з перетином букв, $H^{(k)}$ – ентропія k-грам.

	Алфавіт з пробілом	Алфавіт без пробілу
H_1	4,37484	4,46717
H_2	3,96063	4,14246
H_2^*	3,96241	4,14310

Табл.2.3 – Значення ентропії

$1,5505 < H^{(10)} < 2,4082$
$1,2660 < H^{(20)} < 2,0180$
$1,3224 < H^{(30)} < 2,0779$

Табл.2.4 – Оцінка ентропії

Лабораторная работа №1

Произвольная часть текста:
у_они_перечеркнуто_собственное_утверждение_заявив_что_договор_который_они_со

Использованные буквы:

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

Введенный символ: e

Символ по счету: 1

Номер эксперимента: 60

Неравенство для энтропии:
 $1.55050824161972 < H < 2.40824422811468$

Двоичная таблица угаданных символов:

00100000000000000000000000000000	▲
10000000000000000000000000000000	■
10000000000000000000000000000000	■
00100000000000000000000000000000	■
00010000000000000000000000000000	▼

Вероятности:

q[1]	= 0,5166666
q[2]	= 0,1666666
q[3]	= 0,0833333
q[4]	= 0,05
q[5]	= 0,05
q[6]	= 0,0166666
q[7]	= 0
q[8]	= 0,0333333
q[9]	= 0
q[10]	= 0,0166666
q[11]	= 0,0166666
q[12]	= 0
q[13]	= 0,0166666
q[14]	= 0
q[15]	= 0
q[16]	= 0
q[17]	= 0,0166666
q[18]	= 0
q[19]	= 0
q[20]	= 0
q[21]	= 0
q[22]	= 0
q[23]	= 0
q[24]	= 0,0166666
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 угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

Рис.2.1 – Оцінка $H^{(10)}$

Лабораторная работа №1

Произвольная часть текста:
ди_них_имеется_толь

Использованные буквы:

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

Введенный символ:

Символ по счету:

Номер эксперимента: 61

Неравенство для энтропии:
 $1.26595166494276 < H < 2.01794757892857$

Двоичная таблица угаданных символов:

10000000000000000000000000000000	▲
10000000000000000000000000000000	■
01000000000000000000000000000000	■
10000000000000000000000000000000	■
00001000000000000000000000000000	▼

Вероятности:

q[1]	= 0,6
q[2]	= 0,1666666
q[3]	= 0,0333333
q[4]	= 0,0333333
q[5]	= 0,0666666
q[6]	= 0,0333333
q[7]	= 0
q[8]	= 0
q[9]	= 0
q[10]	= 0
q[11]	= 0
q[12]	= 0,0166666
q[13]	= 0
q[14]	= 0
q[15]	= 0,0166666
q[16]	= 0
q[17]	= 0
q[18]	= 0
q[19]	= 0,0166666
q[20]	= 0
q[21]	= 0,0166666
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:

Рис.2.2 – Оцінка $H^{(20)}$

Лабораторная работа №1

Произвольная часть текста:
таким_же_успехом_можете_предс

Использованные буквы:

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

Введенный символ:

Символ по счету:

Номер эксперимента: 61

Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
1,32242569901381 < H < 2,07788514544078

Двоичная таблица угаданных символов:
00100000000000000000000000000000 ^
10000000000000000000000000000000
00000000000010000000000000000000
00000010000000000000000000000000
10000000000000000000000000000000 v

Вероятности:
q[1] = 0,65
q[2] = 0,0666666
q[3] = 0,0833333
q[4] = 0
q[5] = 0,0166666
q[6] = 0,0166666
q[7] = 0,0166666
q[8] = 0,0166666
q[9] = 0
q[10] = 0,016666
q[11] = 0,016666
q[12] = 0,016666
q[13] = 0,016666
q[14] = 0,033333
q[15] = 0
q[16] = 0
q[17] = 0
q[18] = 0
q[19] = 0
q[20] = 0
q[21] = 0,033333
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:

Рис.2.3 – Оцінка $H^{(30)}$

4. Оцінка надлишковості російської мови

$$R = 1 - \frac{H_{real}}{H_{max}} = 1 - \frac{4.37484}{5} = 0.125$$

ВИСНОВКИ

У даній роботі було обраховано частоти букв і біграм в тексті, а також ентропія за безпосереднім означенням. При підрахунку ентропії біграм як пар літер, що перетинаються, так і пар букв, що не перетинаються, отримані відповіді не сильно відрізнялися одна від одної.

Також було обчислено надлишковість російської мови, отриманий результат перевищує 12%.