КРИПТОГРАФІЯ КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

Експериментальна оцінка ентропії на символ джерела відкритого тексту

Роботу виконали:

студенти групи ФБ-14 Антонова Олександра і Веденкін Артем

Мета роботи

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

Порядок виконання роботи

- 0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.
 - 1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку Н1 та Н2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення Н1 та Н2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення Н1 та Н2 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

Обрали текст: всі томи "Війна і мир" 🤮, автор Лев Толстой, текст написаний російською мовою.

Так як спочатку обраний текст містить багато символів, окрім власне літер, то ми першочергово проводимо фільтрацію тексту та створюємо два файли: з пробілами і без пробілів.

Створюємо функцію для підрахунку частоти букв у тексті, а також обчислення ентропії та надлишку. Результати частоти букв в тексті з пробілами і без пробілів експортуємо у файли «frequency_of_letters_with_spaces.xlsx» і «frequency_of_letters_without_spaces.xlsx» відповідно.

Текст з пробілами

Частота букв:

Α	В	С	D	E	F	G	Н	1	J	K	L	М	N	0	Р	Q	R	S
	a	6	В	г	А	e	ë	э	ж	3	и	ы	й	K	л	м	н	0
частота	0,067302	0,013868	0,040215	0,016975	0,025273	0,067147 0	,000606	0,002434	0,008465	0,014706	0,056073	0,01530	8 0,008783	0,028534	0,040324	0,024707	0,05257	0,094287
Т	- 11		V	W	V	V		7	AA	AB		\C	AD	AE	AF	٨	G	AH
п	p		c	T	y	ф		x	- да - ц	ч		ш	щ	ъ	ь		0	Я
				0,049681		8 0,0017												0,017656

Ентропія:

Надлишковість:

H1: 4.3581476600042395 R: 0.14335538243791024

Текст без пробілів

Частота букв:

Α	В	С	D	E	F	G	Н	1	J	K	L	М	N	0	Р	Q	R
	а	б	В	г	Д	е	ë	э	ж	3	И	ы	й	к	л	M	н
частота	0,08132	0,01676	0,04859	0,02051	0,03054	0,08113	0,00073	0,00294	0,01023	0,01777	0,06775	0,0185	0,01061	0,03448	0,0487	2 0,02985	0,06352
S	T	U	V		W	X	Υ	Z	AA	AB	AC	ΑI) /	AΕ	AF	AG	AH
S	Т	U p	V		W T	X y	У	Z x	AA ц	AB 4	AC w	A[AE ъ	AF b	AG ю	АН я

Ентропія:

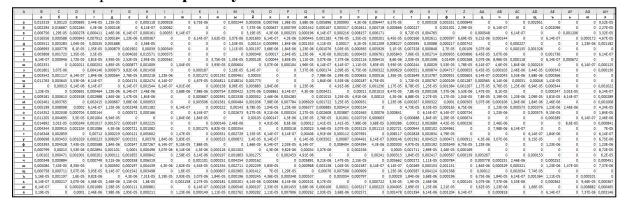
Надлишковість:

H1: 4.4646253219972785 R: 0.11493328705943973

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

Текст з пробілами

Частота біграм без перетину:



Ентропія:

Надлишковість:

H2: 3.976447690967201

R: 0.21838295137504837

Частота біграм з перетином:

a 00						G						M	N	0	P	u	К	5		U	v	W	^		- 4	~~	AD.	ALC	AU	AE	PI	AO	Aff
		6		r	А	e	č		ж	3	И	ы	й	K	л	M	н	0	n	р	c	T	У	ф	x	ц	q	ш	щ	ъ	b	ю	я
				3,38E-05			0,000116			6,45E-06	0	0,000257											0,000339				9,21E-07		0,002947	0	0		9,49E-05
				3,68E-06			3,07E-07			6,14E-07	0									0,004703				0,001302			6,14E-07	0		0,002121	0		2,21E-05
								0,000557		0		3,07E-07											0,004766					6,14E-07					3,41E-05
					0,000195			0		3,07E-07																					6,14E-07		
				0,005356			3,99E-05					1,35E-05														0		0,000231	0	0	3,07E-07		
					0,000875				0		0	9,52E-06																0,000138		0	0		6,14E-07
			9,52E-06				0,015588		0	0	0	0	0,000321	0,000162	2,73E-05	9,21E-07	7,06E-06	0,002185	0,000429	0,006831	0,005895	6,63E-05	0,002746	0,000905	0,000892	6,42E-05	5,22E-06	0	0	0,001742	0		1,53E-06
a 6,	,14E-07 (0,000934	2,06E-05	3,78E-05	3,9E-05	3,01E-05	3,5E-05	0,000556	0	0	3,9E-05	1,53E-05	0,000129	0,000452	5,86E-05	1,14E-05	5,22E-06	7,37E-06	0,002068	0,00642	7,37E-06	2E-05	0,001595	0,014108	0,000272	2,46E-06	8,35E-05	0,000118	0	3,07E-07			9,21E-07
	0,00236				4,02E-05			0	1,26E-05		3,56E-05			2,76E-06	9,98E-05	1,78E-05	3,07E-07	1,23E-06	8,9E-06	6,08E-05	3,96E-05	0,000249	0,000297	6,26E-05	1,75E-05	1,23E-06	4,91E-05	2,76E-06	0,000217	0	0	3,07E-07	
a 0,0	001842	3,68E-06	3,07E-07	0,004133	0,000936	7,67E-06	0,001462	0		0,000185		6,32E-05		0	3,07E-07	8,6E-06	0,00018	2,24E-05	6,75E-06	0,000316	0,000619	0,000557	0,000867	0,004683	0,000635	7,21E-05	0,000104	1,84E-06	3,5E-05	0,000367	0	0	0,000136
					0,000557	2,7E-05	0,003177	9,21E-07	0	0,001265	0,001189	0,000466	0,000346	0	0	0	6,45E-06	2,36E-05	0,000666	0,00052											0	0	0
ы 0,0			6,14E-06				0,002472			5,68E-05				0				6,45E-06										0,000344			0		3,07E-07
й	0 (0,000126	6,75E-06	9,21E-07	3,07E-07	1,23E-06	0,00151	9,21E-07	4,91E-06	0	0,00016	8,84E-05	0,003884	1,84E-06	0	9,21E-07	0	4,91E-06	2,15E-05	0,001239	1,26E-05	9,21E-05	1,29E-05	0,001377	0,001281	1,26E-05	9,76E-05	9,21E-07	6,63E-05	0,000346	0		0,001019
и 1,	,29E-05	0	0,000686	0,000462	1,23E-06	1,53E-06	2,15E-06	0	3,68E-06	1,11E-05	0,000745	0,000422	1,53E-06	0,00682	3,07E-07	5,53E-06	0,003436	0	0,000106	0,00102	7,95E-05	7,77E-05	0,000156	7,98E-06	6,14E-06	1,23E-05	8,53E-05	0	0	0,000254	2,79E-05	0	9,21E-07
			0,003583	0,002984	1,23E-06	1,35E-05	1,53E-06	0	0	0,000696	0,000387	0,002336	0,000345	2,46E-06	0,004249	0	0,000999	3,07E-05	6,14E-06	0,00013											3,07E-07		0,00028
M 0,0	002457 (0,003831	0	0,00239	0,000688	3,07E-06	0,000962	0	0	0,003463	0,001556	0,000497	0,001022	7,98E-06	0,007769	0,00686	0,001702	2,92E-05	0,000933	0	1,53E-06	0,000266	0,000311	0,000102	0,000305	2,76E-06	0,000108	1,53E-06	2,46E-06	2,15E-06	0	0	0,001021
н 0,0	005236	0,008143	0,0002	3,07E-07	6,14E-07	0,00233	0,001172	3,07E-07	3,07E-07	3,07E-07	0,000236	0,001409	8,5E-05	1,14E-05	1,23E-06	0,006315	0,006941	0,000412	0,00013	0	0	4,57E-05	7,98E-05	0,00016	7,37E-06	0	9,21E-07	0,000545	0,003344	1,23E-06	2,46E-06	0	4,18E-05
0 0,0	014372 (0,002565	0,000731	0,000141	0	0,000567	0,000178	6,14E-07	0	3,07E-07	0	0,002685	9,92E-05	8,9E-06	0,000341	2,24E-05	0,00262	0,003915	0,002433	0,000776	0,000814	0	3,07E-07	9,21E-07	0,00073	0	2,15E-06	0	0,000564	9,85E-05	0	0	1,63E-05
n 0,0	011162	0,00473	3,78E-05	0,001007	7,34E-05	0	0	0	3,07E-06	1,78E-05	0	0	0,001029	0,000148	3,38E-06	1,14E-05	2,86E-05	0,009937	0,02066	0,0006	0	0,000876	2E-05	1,26E-05	0,000664	3,07E-07	0	0	0	0,000281	0	3,07E-07	1,84E-06
p 0,0	014847	2,64E-05	0,000259	0,001002	0,001599	0,000177	0,002239	0	0	0	0,000148	2,76E-06	0	4,91E-06	7,37E-06	0,000117	1,57E-05	1,41E-05	8,9E-06	3,71E-05	0,000287	0,000127	0,002654	4,51E-05	0,00025	0	1,53E-06	0	0	0	8,47E-05	0	0,000843
c 0,0	004366 0	0,000892	0,002096	0,001087	3,38E-06	0,000741	0,00106	0	0	0,001264	6,85E-05	0,000366	1,23E-06	0	0,000516	0,000235	5,56E-05	2,76E-05	0,003118	0,001128	0,002734	0,000941	0,000107	0,004982	0	0	7,67E-06	3,07E-07	0	0	0	3,07E-07	7,15E-05
T 0,0	016389 (0,002823	0	0,00713	0,000232	0,000119	0,005608	0	1,07E-05	0	0,000335	0,002743	1,93E-05	1,84E-06	9,21E-07	0,004058	5,83E-06	0,000117	0,007655	6,14E-07	0,008064	0,002757	0,00284	4,51E-05	0	0	0	0	9,21E-07	1,53E-06	0	0	6,14E-07
y 0	0,00671 (0,003463	0,001067	0,000226	0,008364	0,001547	0,00371	1,53E-06	4,91E-05	0	0,00053	0,000876	5,28E-05	3,07E-07	0	4,3E-06	0,003528	0,000883	0,004466	3,07E-07	0,000153	0,00138	0,004466	0,000488	0,000912	3,07E-06	2,46E-06	0	8,53E-05	0	0	0	4,6E-06
ф 0,0	003437 (0,004532	7,18E-05	0,000601	1,84E-06	0,003459	0,007281	6,14E-07	8,81E-05	7,67E-05	0	0	1,78E-05	1,53E-06	1,96E-05	3,07E-07	3,07E-07	0,008355	0,004416	3,68E-06	0,000286	5,03E-05	0,00517	0,005542	6,51E-05	2,15E-06	0	0	0	1,84E-06	0	0	1,23E-06
x 0,0	.000827 (0,000131	6,45E-06	0,002908	0,00112	9,49E-05	0,000516	1,84E-06	4,6E-06	3,99E-06	0,000116	0,001299	0	4,91E-06	9,52E-06	0,000338	2,15E-06	3,07E-07	0,002243	0	0,000305	0,001705	2,76E-05	2,21E-05	0,000171	0	0	0	0	0	0	0	8,07E-05
ц 0,7	001032 0	0,000483	0,000967	0,000205	0,000107	0,001795	0,000857	0	2,7E-05	6,45E-06	0,000196	0,001848	0,001323	0	0,002459	4,91E-06	3,07E-07	2,15E-06	0	0,002396	0,000318	1,93E-05	0,002445	0,000544	0,000137	0,003263	0	0	0,000165	0	0	0	5,53E-05
4 0,0	000323 (0,000908	0	0,000768	8,29E-06	0,000375	0,006302	0	0	0,001092	0,000247	0,004307	0,000173	3,07E-07	0	0,005849	6,14E-06	1,53E-05	2,3E-05	0	0,00571	0,000281	1,26E-05	0,000792	3,07E-07	0,000774	0,000225	3,07E-06	0	0,000244	0	0	0,000431
ш 0	0,00512	7,15E-05	8,9E-05	0	0,000602	0,000195	0,004431	5,53E-06	2,73E-05	4,85E-05	0,001524	0,00277	0,0022	6,32E-05	0	0,000942	0,003094	1,35E-05	0,001609	6,14E-07	3,07E-07	0,002647	0,00015	0,00113	0	9,21E-07	0,000522	0,000318	0	6,14E-07	1,41E-05	0	7,98E-06
щ 0,0	000727	0,00072	2,76E-06	3,71E-05	3,07E-07	0,00153	0,003553	9,21E-07	1,69E-05	0	0,000619	0,002992	0,001395	7,55E-05	1,63E-05	0	0,000755	0,007596	0,000898	0	9,21E-07	0,000588	0,004151	0,001589	3,07E-07	0,000118	0	0,002021	8,2E-05	0	0	0	0
ъ 5,	,62E-05 (0,001197	1,78E-05	1,17E-05	0	5,53E-06	6,97E-05	3,53E-05	3,84E-05	2,76E-06	1,93E-05	0,000194	0,000247	4,08E-05	0,000362	0,000534	0	0,003061	0,000777	0	0,000303	9,21E-07	3,68E-06	0,000196	0	6,75E-06	2,06E-05	3,07E-07	0,001353	9,52E-06	0	0,000102	0
ь 3,	07E-07 (0,000228	3,38E-06	3,96E-05	1,84E-06	3,9E-05	1,5E-05	0	0,002139	2,21E-05	0,000195	0,000206	6,45E-06	0,000401	5,83E-06	9,7E-05	7,28E-05	0	0	0,000737	3,19E-05	1,96E-05	2,15E-06	0	0,000133	3,99E-06	7,37E-06	4,91E-06	0	0,000355	9,21E-07	9,06E-05	0,000375
ю 9,	21E-07	0	0,00019	0,001096	2,46E-05	0,000111	0,000866	0	. 0	3,07E-07	0,000214	0,000525	0,00011	1,93E-05	0,001438	7,67E-06	0,000103	0,000202	0,003187	0,000238	0,004842	2,64E-05	9,21E-07	2,09E-05	0	3,07E-05	9,21E-07	3,07E-07	1,66E-05	0	0	0,008942	0,000455
я 3.	.47E-05	0	0.000101	3.99E-06	7.06E-06	3.1E-05	0.00022	0	1.53E-06	0.000144	1.41E-05	0.002854	0.000276	1.32E-05	0.007851	0.000266	2.43E-05	4.91E-06	0.002567	0	0.001451	0.001368	6.45E-06	0.001187	3.07E-07	0	0.000848	0	0	3.07E-07	0	6.14E-06	0.003075

Ентропія:

Надлишковість:

H2: 3.9770017429061086

R: 0.21827404602159495

Текст без пробілів

Частота біграм без перетину:

A	В	С	D	E	F	G	Н	- 1	J	K	L	M	N	0	P	Q	R	\$	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
	a	6		r	A	e	ě		ж	3	и	M	й	K	л	M	н	0	n	р	c	T	y	Φ	×	ц	4	w	щ	ъ	, b	ю	Я
a	0,000268	0,00116	0,006604	0,001059	0,004802	0,000217	3,71E-06	0	0,001513	0,00579	0,00032	5,93E-05	0,0002	0,009499	0,007697	0,003271	0,012011	0,0002	0,001352	0,009856	0,001737	0,006287	0,000142	0,000197	0,000984	0,000621	0,002418	0,001672	0,000423	0	0,000438	0,000124	0,000299
6	0,001482	1,48E-05	0,000256	7,86E-05	7,72E-05	0,002157	6,16E-05	7,42E-07	8,23E-05	0,000227	0,001441	0,000662	0,000341	0,000464	0,000314	0,000332	0,00032	0,005115	7,42E-07	0,000251	0,000189	0,000302	0,000872	5,93E-06	0,000145	7,42E-06	1,19E-05	1,19E-05	7,42E-07	0	0,000438	0,00055	0,000534
8	0,006733	0,000116	0,000576	0,000287	0,001165	0,003608	6,16E-05	2,97E-06	2,23E-05	0,000993	0,005097	0,001645	0,000779	0,000781	0,000827	0,000818	0,000557	0,01208	4,45E-06	0,00095	0,002235	0,003554	0,001564	2,67E-05	0,000524	5,19E-05	4,3E-05	2,23E-05	1,48E-06	0	0,001227	0,000326	0,001651
r	0,001484	1,26E-05	0,000489	8,46E-05	0,001102	0,00439	1,19E-05	7,42E-06	2E-05	0,000519	0,001218	0,000215	0,00034	0,000145	0,000384	0,000381	0,000289	0,005886	5,19E-06	0,000385	0,000146	0,00016	0,001543	2,97E-06	0,000131	1,93E-05	0,000155	3,71E-06	0	0	0,000272	0,000139	0,000495
А	0,002979	2,3E-05	0,000734	0,001271	8,46E-05	0,003893	2,82E-05	8,16E-06	0,000893	0,001959	0,002729	0,000358	0,000684	0,000243	0,000628	0,000393	0,001199	0,005424	2,23E-06	0,000442	0,00052	0,000458	0,002248	3,71E-06	0,000208	9,64E-06	5,19E-06	7,42E-06	0	0	0,000436	0,000662	0,001202
e	0,001698	0,002398	0,005188	0,000698	0,005022	0,002208	2,37E-05	7,42E-07	0,004217	0,000298	0,003771	0,001339	0,000123	0,0007	0,005462	0,004445	0,010187	0,003341	0,002946	0,00693	0,003217	0,005332	0,000347	0,000151	0,000101	0,001023	0,00352	0,002548	0,001451	0,000209	0,001921	7,27E-05	0,000483
ě	0	7,42E-07	1,48E-06	0	2,23E-06	3,71E-06	0	0	7,42E-07	0	0	0	0	7,42E-07	5,93E-06	1,48E-06	7,42E-07	0	0	0	0,000701	4,45E-06	0	0	0	0	0	0	7,42E-07	7,42E-07	7,42E-07	0	0
	0,000346	4,3E-05	0,0003	3,41E-05	2,3E-05	0,000243	7,64E-05	1,48E-06	1,34E-05	4,01E-05	0,000254	5,79E-05	6,9E-05	9,35E-05	0,000116	7,49E-05	4,45E-05	0,000381	1,48E-06	2,08E-05	5,71E-05	0,000128	9,5E-05	0	3,26E-05	5,19E-06	2,97E-06	1,48E-06	0	0	0,000203	4,3E-05	0,000151
ж	0,00151	1,04E-05	6,01E-05	1,56E-05	2,37E-05	0,001305	6,68E-06	0	2,37E-05	0,000194	0,000639	7,12E-05	0,000137	0,000201	0,000453	0,000124	5,49E-05	0,00238	7,42E-07	0,000375	0,00012	0,000116	0,001509	0	4,9E-05	1,48E-06	2,23E-06	7,42E-07	0	0	8,38E-05	9,27E-05	0,000657
3	0,004823	8,9E-06	0,000698	6,38E-05	5,64E-05	0,001932	2,3E-05	3,71E-06	5,93E-06	0,000119	0,003861	0,000222	0,000151	9,42E-05	0,000157	0,000166	0,000177	0,001616	1,48E-06	7,27E-05	0,000106	0,000137	0,001151	5,19E-06	8,68E-05	5,93E-06	2,97E-06	7,42E-07	0	0	0,000418	0,0001	0,001474
и	0,001493	0,000871	0,004969	0,001006	0,002927	0,001495	1,26E-05	0	0,00149	0,000479	0,002736	0,000497	0,000623	0,003345	0,007436	0,004548	0,009453	0,002413	0,000911	0,005866	0,001778	0,004375	0,00065	0,000694	0,000518	0,000328	0,001642	0,002109	0,000829	7,42E-07	0,001012	0,000326	0,000983
ы	0	0,00431	0,002875	1,48E-06	0,000663	0	0	0	0	0,00064	0	0	0	0	0,001155	0,000901	0,003676	0	0,000282	0,001764	0,00037	0,001567	0	1,56E-05	0	0,000176	0	0	0	0	0	0	0
й	0,000754	0	0	7,42E-07	7,42E-07	0,002737	0	1,19E-05	0	0	0,00159	0,001608	2,23E-06	0	7,42E-07	0	0	0,003734	0	0	0	7,42E-07	7,49E-05	0	7,42E-07	0	0	0	0	0	7,42E-07	0	5,04E-05
K	0,005907	0,000235	0,000835	0,000209	0,000292	0,002211	2,45E-05	6,31E-05	0,000116	0,000275	0,003385	0,000436	0,000674	0,000399	0,001295	0,000618	0,000446	0,004209	8,16E-05	0,000368	0,005972	0,000731	0,001355	1,34E-05	0,000207	0,00016	0,000266	0,000469	2,23E-06	0	0,001934	0,000253	0,001005
А	0,010116	0,000896	0,001373	0,001963	0,000902	0,007104	8,16E-06	9,87E-05	8,9E-06	0,000284	0,005516	0,002719	0,000239	0,000474	0,000201	0,000305	4,38E-05	0,008052	0,000849	0,000128	0,003203	0,000343	0,001456	7,49E-05	0,000286	5,93E-06	4,6E-05	0,000464	0	0	0,000144	0,000131	0,00112
M	0,003828	4,97E-05	0,001556	5,04E-05	0,000188	0,005508	9,64E-06	1,11E-05	2,23E-05	0,000368	0,003952	0,001909	0,000378	0,000189	0,000142	0,000261	0,000127	0,006418	2,23E-06	0,000397	0,001171	0,000212	0,001297	1,04E-05	0,000198	1,046-05	1,34E-05	2,15E-05	0	0	0,000609	0,000154	0,000723
н	0,007748	0,00032	0,001875	0,00045	0,002021	0,010304	2,6E-05	3,19E-05	0,00135	0,001942	0,00558	0,00082	0,00108	0,002252	0,001057	0,001863	0,003676	0,009816	8,23E-05	0,000878	0,001436	0,001601	0,001091	1,71E-05	0,000412	3,12E-05	0,000605	0,000408	4,15E-05	0	0,002188	0,00031	0,002167
0	0,001453	0,002669	0,009245	0,010343	0,004222	0,001887	3,12E-05	7,42E-07	7,42E-05	0,000813	0,001876	0,000309	0,000524	0,009893	0,008031	0,004579	0,010293	0,00236	0,0099	0,009097	0,003483	0,017504	0,000426	0,000114	0,00271	0,000423	9,94E-05	0,000322	0	0	0,000768	0,000232	0,000803
п	0,002987	8,9E-06	0,001108	0,000259	0,000217	0,002932	4,53E-05	3,93E-05	1,41E-05	0,000214	0,002576	0,000594	0,000799	0,000389	0,000925	0,001023	0,000593	0,003738	5,42E-05	0,000286	0,002223	0,000523	0,001358	1,48E-05	0,000447	4,08E-05	2,97E-05	5,64E-05	0	0	0,000973	0,000319	0,001082
р	0,003933	0,001287	0,001436	0,001342	0,002211	0,008098	4,67E-05	2,23E-05	5,93E-06	0,000287	0,001217	0,00045	0,000331	0,001804	0,000252	0,000249	0,000171	0,007717	0,007142	5,04E-05	0,000452	0,005775	0,000878	0,000706	0,000246	7,42E-06	4,01E-05	2,97E-06	5,93E-06	0	0,000236	0,000209	0,000323
c	0,006129	8,83E-05	0,00453	0,000301	0,000581	0,007455	5,42E-05	5,12E-05	4,45E-05	0,000447	0,005568	0,001362	0,001706	0,000686	0,002238	0,001099	0,001174	0,010405	1,41E-05	0,001787	0,001387	0,001595	0,002194	4,01E-05	0,000545	2,82E-05	2,3E-05	5,93E-06	0	7,42E-07	0,002051	0,000539	0,002053
T	0,00625	9,64E-06	0,000786	6,53E-05	0,000266	0,004973	0,000104	0,00256	1,85E-05	8,46E-05	0,004555	0,000956	0,000535	0,000769	0,000338	0,000317	0,000725	0,008938	4,82E-05	0,000708	0,01687	0,000408	0,001778	2,74E-05	0,000162	1,85E-05	0,004008	0,000131	0	0	0,000507	0,000618	0,002091
Y	0,000562	0,001183	0,001082	0,000787	0,001901	0,000573	1,48E-05	0	0,000176	0,000341	0,00052	0,000102	0,000163	0,001729	0,001473	0,002801	0,002651	0,000622	0,000733	0,003274	0,001066	0,001795	0,000117	0,0001	0,000235	0,000684	0,000689	0,000348	0,000106	7,42E-07	0,000195	0,000124	0,000233
Φ	0,000635	0	3,56E-05	1,41E-05	1,93E-05	9,79E-05	7,42E-07	1,48E-06	7,42E-07	7,42E-06	0,000158	2,45E-05	7,57E-05	2,67E-05	4,53E-05	6,01E-05	3,12E-05	0,000494	1,85E-05	3,64E-05	5,49E-05	3,86E-05	4,38E-05	6,68E-06	2,3E-05	2,23E-06	1,48E-06	7,42E-07	0	0	3,93E-05	2,6E-05	5,64E-05
x	0,001205	0,000112	7,72E-05	1,19E-05	3,71E-05	0,001214	7,42E-06	8,16E-06	1,48E-06	5,19E-06	0,002148	0,001245	6,23E-05	4,9E-05	3,04E-05	6,45E-05	5,27E-05	0,000805	7,42E-07	0,000129	0,000208	5,27E-05	0,000403	1,48E-06	2,08E-05	7,42E-07	1,48E-06	7,42E-07	0	0	7,86E-05	2,97E-05	0,0003
ц	0,000109	2,97E-06	4,15E-05	6,68E-06	0,00015	0,000295	7,42E-07	1,48E-06	0	8,16E-06	0,001425	1,48E-05	7,27E-05	2,3E-05	9,64E-06	4,01E-05	0,000779	0,000111	1,56E-05	9,05E-05	3,93E-05	0,000128	2,37E-05	7,42E-07	1,11E-05	0	0	2,23E-06	0	0	0,000102	5,04E-05	7,12E-05
q	0,001353	2,52E-05	0,000228	0,000156	7,05E-05	0,001571	5,56E-05	1,48E-06	3,04E-05	3,86E-05	0,002088	0,000293	0,000303	0,000154	0,000607	0,000503	0,000352	0,002709	9,42E-05	0,000119	0,000431	0,00042	0,001306	1,48E-06	8,23E-05	5,19E-06	6,68E-06	1,48E-06	0	0	0,000447	0,000216	0,000504
w	0,00152	3,71E-06	0.001402	1.19E-05	0,000152	0,000817	0	7,42E-07	0	2,67E-05	0,000518	0,00058	0.000197	2,74E-05	4,67E-05	6,97E-05	4,3E-05	0,001327	5.19E-06	0,000323	0,000122	1,34E-05	0,000816	2,97E-06	2,67E-05	1,48E-05	0.000103	3,71E-06	0	0	0.000522	0,000108	6,45E-05
щ	0,000294	0,000222	1,11E-05	7,42E-07	7,42E-07	0,00066	0	0	0	2,23E-06	0,000151	5,19E-06	1,48E-06	2,97E-06	6,68E-06	6,68E-06	0,000119	0,000154	2,97E-06	5,64E-05	0	1,85E-05	0,000321	0	1,48E-06	0	0	0	0	0	1,34E-05	0,000562	0,000262
ъ	0	0,000127	1,63E-05	0	0,000177	0	0	0	0	1,41E-05	7,42E-07	0	0	0	0	0	8,16E-06	0	0	0	2E-05	5,34E-05	0	0	2,97E-06	0	0	0	0	0	0	0	0
ь	0	4,38E-05	0,000187	7,42E-07	0,000434	0	0	0	4,53E-05	0,000715	0	0	0	1,48E-06	0,004218	3,41E-05	0,001121	0	0,000984	0,001045	0,003217	0,0061	7,42E-07	2,97E-06	0	0	0,000191	0,000304	1,78E-05	0	0	0	0
ю	0,000866	5,93E-06	3,71E-06	2,97E-06	5,49E-05	0,000424	0	0	4,23E-05	8,09E-05	0,000435	7,42E-07	5,19E-06	6,68E-06	0,001223	3,93E-05	0,000162	0,000912	0	0,000157	0,000131	6,08E-05	0,001097	0	7,42E-07	0	0	7,42E-07	0	0,000137	0,000482	4,15E-05	0,000128
я	0,002947	0,000506	0,000266	1,26E-05	0,000492	0,000582	7,42E-07	0	6,688-06	0,000585	0,002363	4,08E-05	6,23E-05	6,08E-05	0,002115	0,000535	0,003019	0,001008	0,000357	0,001134	0,003486	0,000637	0,000142	2,97E-06	2,6E-05	2,97E-06	7,42E-07	0	0	0,000111	0,000634	4,01E-05	0,000263
																																_	

Ентропія:

Надлишковість:

H2: 4.145533075985768

R: 0.1781900902475484

Частота біграм з перетином:

A	В	C	D	Ε	F	G	H	1	1	K	L	M	N	0	P	Q	R	\$	T	U	٧	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH
	a	6	5	r	A	e	ë	,	NK.	3	и	ы	й	K	л	M	н	0	n	р	c		y	•	x	ц	q	ш	щ	ъ	ь	10	я
a	0,000277	0,001129	0,006592	0,001082	0,004891	0,000211	2,6E-06	3,71E-07	0,00153	0,005803	0,000336	5,3E-05	0,000184	0,009451	0,00774	0,003271	0,012054	0,000206	0,001364	0,009794	0,001708	0,006323	0,000146	0,000171	0,000977	0,000635	0,002446	0,001637	0,000429	0	0,000463	0,000119	0,000298
6	0,001468	1,6E-05	0,00027	7,98E-05	7,98E-05	0,00218	5,9E-05	3,71E-07	7,79E-05	0,000233	0,001424	0,000669	0,00033	0,000453	0,000302	0,000342	0,00032	0,005109	1,48E-06	0,000246	0,000194	0,000314	0,000871	9,27E-06	0,000155	1E-05	1,15E-05	1,04E-05	3,71E-07	0	0,000447	0,000563	0,00051
0	0,006791	0,000106	0,000583	0,000287	0,001179	0,003638	5,75E-05	3,71E-06	2,23E-05	0,00101	0,005142	0,001676	0,000774	0,000783	0,000815	0,00081	0,00055	0,012089	3,71E-06	0,000959	0,002269	0,003565	0,001545	2,15E-05	0,000519	5,27E-05	4,08E-05	2,78E-05	1,85E-06	3,71E-07	0,001275	0,000324	0,001667
r	0,001451	1,34E-05	0,000492	8,49E-05	0,001096	0,004362	1,11E-05	5,93E-06	1,89E-05	0,000512	0,001237	0,000208	0,000332	0,000152	0,000379	0,000379	0,000277	0,005988	5,93E-06	0,0004	0,000156	0,000155	0,00155	4,45E-06	0,000146	1,89E-05	0,000149	2,97E-06	0	0	0,000279	0,000137	0,000509
								7,05E-06																								0,000662	
e	0,001686	0,002424	0,005172	0,0007	0,005076	0,002197	1,97E-05	3,71E-07	0,004193	0,000306	0,003752	0,001347	0,000117	0,000701	0,005355	0,004395	0,010167	0,003333	0,002895	0,006916	0,003247	0,005221	0,000344	0,000162	0,0001	0,001032	0,003563	0,002564	0,001444	0,000205	0,001936	8,38E-05	0,000477
ě		3,71E-07			1,48E-06		0		3,71E-07			0					3,71E-07				0,000711		0	0	0	0			7,42E-07	7,42E-07	7,42E-07	0	C
9	0,000328	4,08E-05	0,000298	3,12E-05	2,6E-05	0,000246	6,83E-05	7,42E-07	1,26E-05	4,15E-05	0,000252	5,49E-05	6,75E-05	0,000103	0,000116	7,72E-05	4,75E-05	0,000378	2,23E-06	2,19E-05	5,82E-05	0,000119	0,000101	3,71E-07	3,89E-05	4,82E-06	2,6E-06	2,97E-06	3,71E-07	0	0,000197	4,3E-05	0,000161
ж	0,00149	8,53E-06	5,9E-05	1,74E-05	2,41E-05	0,001378	4,82E-06	0	2,26E-05	0,000189	0,000629	6,27E-05	0,000133	0,000213	0,000453	0,000127	5,3E-05	0,002348	3,71E-07	0,000375	0,000123	0,00011	0,00151	1,85E-06	4,71E-05	1,48E-06	1,85E-06	1,48E-06	0	0	7,01E-05	8,9E-05	0,000684
3	0,004842	1,15E-05	0,0007	6,31E-05	4,97E-05	0,001988	2,41E-05	3,71E-06	8,16E-06	0,000115	0,003861	0,000205	0,000156	9,72E-05	0,000151	0,000169	0,000175	0,001621	7,42E-07	7,86E-05	0,000103	0,000123	0,001158	5,19E-06	8,2E-05	5,56E-06	3,71E-06	2,23E-06	7,42E-07	0	0,000409	9,2E-05	0,001462
и	0,001525	0,000871	0,004924	0,00098	0,002921	0,00151	1,52E-05	3,71E-07	0,001538	0,00045	0,002724	0,000477	0,000616	0,003284	0,007529	0,004478	0,009521	0,002425	0,000897	0,005979	0,001761	0,004321	0,00065	0,000674	0,000514	0,000303	0,001633	0,002111	0,000837	3,71E-07	0,000987	0,000335	0,000961
ы		0,004329	0,002888	7,42E-07	0,000685	0	0	0	0	0,000641	0	0	0	0	0,001138	0,000912	0,003698		0,000287	0,001753	0,00041	0,001573	0	1,22E-05	0	0,000167	0	3,71E-07	0	0	0	0	C
й	0,000739	0	3,71E-07	3,71E-07	3,71E-07	0,002707	0	1,3E-05	0	0	0,001571	0,001599	2,6E-06	3,71E-07	1,48E-06	7,42E-07	0	0,003852	0	3,71E-07	0	7,42E-07	7,75E-05	0	3,71E-07	0	0	0	0	0	1,11E-06	0	4,45E-05
к	0,005886	0,000244	0,000833	0,000197	0,000312	0,002191	2,3E-05	5,97E-05	0,000103	0,000276	0,003394	0,000434	0,00066	0,000372	0,001277	0,000636	0,000443	0,004279	8,09E-05	0,000364	0,005991	0,000743	0,001395	1,26E-05	0,00023	0,00016	0,000273	0,00045	1,48E-06	3,71E-07	0,001916	0,000251	0,00099
Л	0,010124	0,000883	0,001381	0,001948	0,000911	0,007152	1,11E-05	0,000107	8,16E-06	0,000262	0,005543	0,002712	0,000236	0,000492	0,000189	0,000294	4,56E-05	0,008125	0,000824	0,000116	0,003238	0,000352	0,001465	8,01E-05	0,000287	6,31E-06	4,53E-05	0,000476	0	0	0,000154	0,000133	0,001122
M	0,003847	4,86E-05	0,001523	4,67E-05	0,00018	0,005499	9,64E-06	7,05E-06	2,19E-05	0,00038	0,003975	0,001899	0,000393	0,000193	0,000137	0,000261	0,000115	0,00645	1,85E-06	0,000389	0,00117	0,000202	0,00134	7,42E-06	0,000197	1,11E-05	1,04E-05	2,23E-05	3,71E-07	0	0,000618	0,000162	0,000736
н	0,00773	0,000323	0,001888	0,00047	0,002012	0,010281	2,89E-05	3,3E-05	0,001347	0,001954	0,005641	0,00083	0,001056	0,002286	0,001052	0,001836	0,003638	0,009855	8,23E-05	0,000881	0,001428	0,001588	0,001061	1,48E-05	0,000412	2,82E-05	0,000608	0,000425	3,45E-05	0	0,002209	0,00033	0,002157
0	0,001415	0,002644	0,009113	0,010254	0,004246	0,001914	3,45E-05	3,71E-07	7,01E-05	0,000832	0,001868	0,000301	0,000527	0,009871	0,007982	0,004572	0,010192	0,00233	0,009964	0,009026	0,003503	0,017402	0,000414	0,000104	0,002689	0,000435	9,83E-05	0,000313	3,71E-07	3,71E-07	0,000759	0,000242	0,00080€
n	0,002969	7,42E-06	0,001102	0,000256	0,000233	0,00297	4,82E-05	3,49E-05	1,3E-05	0,000206	0,002559	0,000606	0,000794	0,000389	0,000962	0,001043	0,000573	0,003719	5,9E-05	0,000274	0,002215	0,000509	0,001362	2E-05	0,00044	3,49E-05	3,38E-05	5,45E-05	7,42E-07	0	0,000965	0,000316	0,001112
р	0,003908	0,001292	0,001406	0,001397	0,002209	0,008087	4,12E-05	2,04E-05	5,19E-06	0,000284	0,001243	0,000445	0,000352	0,001812	0,00024	0,000248	0,000182	0,007742	0,007122	5,97E-05	0,000467	0,005768	0,000875	0,00068	0,000238	7,42E-06	4,12E-05	7,05E-06	4,08E-06	0	0,00023	0,000209	0,000332
c	0,006055	9,53E-05	0,004508	0,000324	0,000599	0,007404	5,68E-05	4,71E-05	4,15E-05	0,000431	0,005554	0,001346	0,001713	0,000695	0,00224	0,001077	0,001183	0,010528	1,37E-05	0,001768	0,001395	0,001579	0,002194	3,56E-05	0,000533	3,56E-05	2,82E-05	8,16E-06	7,42E-07	3,71E-07	0,002051	0,000511	0,002116
T	0,006307	9,64E-06	0,000808	6,23E-05	0,000274	0,005045	0,000104	0,002584	1,89E-05	8,75E-05	0,004615	0,00096	0,000527	0,000782	0,000347	0,000318	0,000698	0,008978	4,93E-05	0,00072	0,016852	0,000405	0,001807	2,26E-05	0,000159	1,48E-05	0,004047	0,000121	3,71E-07	0	0,000543	0,000649	0,002118
y	0,000555	0,001204	0,001065	0,000772	0,001872	0,000581	1,56E-05	0	0,000177	0,000339	0,000521	9,64E-05	0,000174	0,001752	0,001483	0,002834	0,00267	0,000618	0,000748	0,003278	0,001025	0,001764	0,000121	9,46E-05	0,000231	0,00066	0,000688	0,00034	0,000102	3,71E-07	0,000203	0,000117	0,00024
ф	0,000693	0	3,34E-05	1,11E-05	2,23E-05	9,24E-05	7,42E-07	1,85E-06	3,71E-07	7,79E-06	0,000173	2,63E-05	7,68E-05	2,49E-05	4,75E-05	6,05E-05	3,12E-05	0,000503	2,26E-05	3,82E-05	4,71E-05	3,49E-05	4,64E-05	4,82E-06	2,71E-05	1,48E-06	7,42E-07	3,71E-07	3,71E-07	0	4,04E-05	2,45E-05	6,31E-05
X	0,001205	0,000108	8,68E-05	9,64E-06	5,08E-05	0,001219	7,79E-06	7,79E-06	1,11E-06	6,31E-06	0,00217	0,001235	6,83E-05	4,08E-05	3,89E-05	7,05E-05	5,64E-05	0,000792	3,71E-07	0,000123	0,000203	5,04E-05	0,000413	1,48E-06	1,85E-05	1,48E-06	2,23E-06	7,42E-07	0	0	8,05E-05	2,74E-05	0,000296
ц	0,000136	3,34E-06	4,19E-05	5,19E-06	0,000138	0,000296	1,11E-06	1,11E-06	0	8,53E-06	0,001382	1,52E-05	6,38E-05	2,52E-05	1,08E-05	4,04E-05	0,00081	0,000123	1,52E-05	9,61E-05	4,04E-05	0,000134	1,97E-05	3,71E-07	1,19E-05	0	0	2,6E-06	0	0	0,000111	4,49E-05	7,01E-05
4	0,001342	2,41E-05	0,000226	0,000165	6,9E-05	0,001572	5,53E-05	1,11E-06	4,6E-05	4,01E-05	0,00204	0,000291	0,000317	0,000155	0,000593	0,000495	0,000343	0,002691	9,76E-05	0,000109	0,000413	0,000398	0,00129	2,23E-06	7,57E-05	5,56E-06	7,79E-06	1,85E-06	0	0	0,000458	0,000239	0,000486
w	0,001528	4,82E-06	0,001379	1,26E-05	0,000147	0,000838	2,23E-06	3,71E-07	3,71E-07	2,34E-05	0,000521	0,000564	0,00019	3,08E-05	5,19E-05	6,64E-05	4,19E-05	0,001316	6,31E-06	0,00033	0,000116	1,68-05	0,000824	2,23E-06	2,78E-05	1,48E-06	0,000105	2,6E-06	0	0	0,000556	0,000119	6,38E-05
щ	0,000283	0,00023	8,9E-06	3,71E-07	7,42E-07	0,000676	0	0	0	1,11E-06	0,000166	5,93E-06	2,97E-06	2,23E-06	6,31E-06	8,53E-06	0,00013	0,000163	2,23E-06	5,53E-05	1,48E-06	1,78E-05	0,000306	0	1,48E-06	0	0	0	0	0	1,15E-05	0,000548	0,000249
ъ	0	0,000122	1,45E-05	0	0,000196	0	0	0	0	1,85E-05	3,71E-07	0	0	0	0	0	7,42E-06	0	0	0	2,41E-05	5,45E-05	0	0	2,97E-06	0	0	0	0	0	0	0	
ь	0	4,08E-05	0,00018	1,11E-06	0,000437	0	0	0	4,71E-05	0,00071	7,42E-07	0	0	1,11E-06	0,004152	3,71E-05	0,001127	3,71E-07	0,000984	0,001058	0,003207	0,006019	3,71E-07	3,71E-06	0	0	0,0002	0,000295	1,71E-05	0	0	0	(
ю	0,000848	4,45E-06	4,82E-06	3,34E-06	4,93E-05	0,000435	0	0	4,3E-05	7,64E-05	0,000424	3,71E-07	4,82E-06	8,53E-06	0,001208	3,64E-05	0,000158	0,000941	0	0,000154	0,000131	5,53E-05	0,001103	2,6E-06	7,42E-07	0	3,71E-07	7,42E-07	0	0,000123	0,000456	4,08E-05	0,000136
я	0,002963	0,000491	0,000268	1,26E-05	0,000474	0,000576	1,11E-06	0	6,688-06	0,000567	0,002345	4,15E-05	6,49E-05	6,34E-05	0,002097	0,000562	0,002965	0,001014	0,000348	0,001154	0,003443	0,000654	0,000139	2,6E-06	2,49E-05	4,45E-06	1,85E-06	7,42E-07	0	0,000109	0,000647	3,64E-05	0,000254

Ентропія:

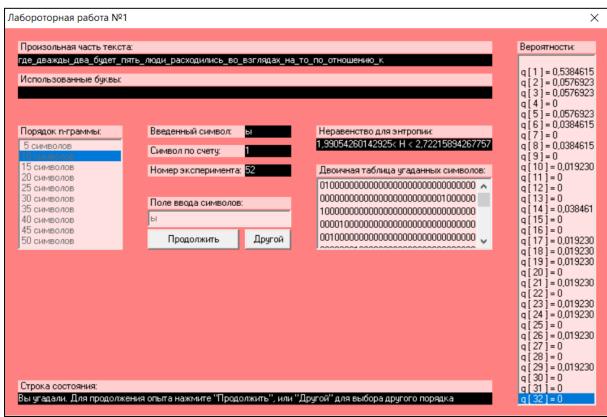
Надлишковість:

H2: 4.145662874396046

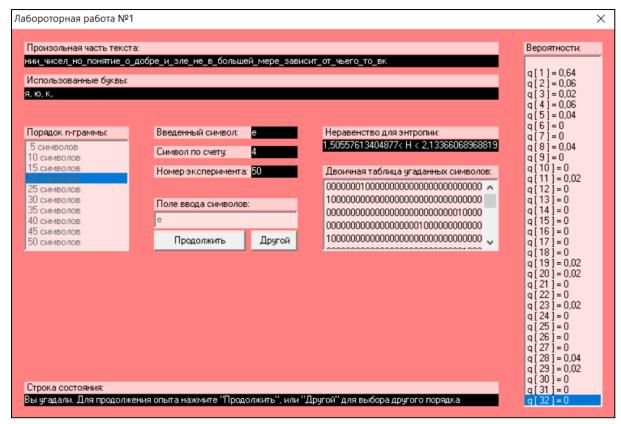
R: 0.17816435902845518

Використовуючи програму CoolPinkProgram, оцінимо значення $H^{(10)}$ $H^{(20)}$, $H^{(30)}$:

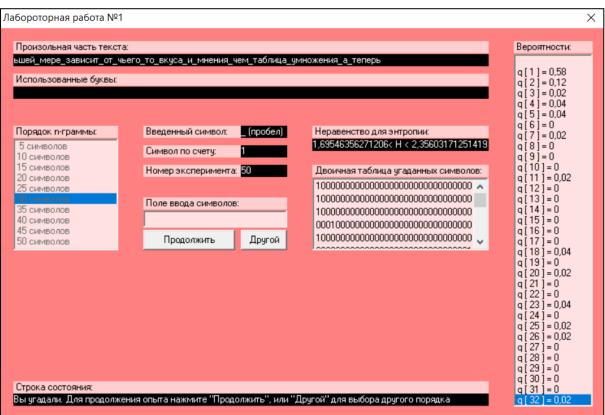
$1.99054260142925 < H^{(10)} < 2.72215894267757$



$1.50557613404877 < H^{(20)} < 2.13366068968819$



$1.69546356271206 < H^{(30)} < 2.35603171251419$



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

- для H⁽¹⁰⁾ :
 - 1. R = 0.6018914791415
 - 2. R = 0.455568211464486
- для H⁽²⁰⁾ :
 - 1. R = 0.698884773190246
 - 2. R = 0.573267862062362
- для H⁽³⁰⁾:
 - 1. R = 0.660907287457588
 - 2. R = 0.528793657497162

Надлишковість обчислюємо за формулою $R = 1 - \frac{H_{\infty}}{H_0}$, $H_0 = log_2 32 = 5$.

Висновки:

Під час виконання комп'ютерного практикуму ми ознайомились з поняттям ентропії на символ джерела та його надлишковості, написали програми для підрахунку частот букв і частот біграм в тексті, попрацювали з CoolPinkProgram, де оцінювали та порівнювали $H^{(10)}$, $H^{(20)}$, $H^{(30)}$. В результаті експериментів ми отримали значення ентропії та надлишковості

власного джерела, які наближаються до значень, отриманих з прикладеної програми.