НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО»
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

МЕТОДИ КРИПТОАНАЛІЗУ 1

КОМП'ЮТЕРНИЙ ПРАКТИКУМ №2

Статистичні критерії на відкритий текст

Варіант 3

Виконав:
Бєш Радомир ФІ-42мн
Перевірив:
Ядуха Д.В.

Київ — 2025

Зміст

1 Мета

Засвоєння статистичних методів розрізнення змістовного тексту від випадкової послідовності, порівняння їх, визначення похибок першого та другого роду.

2 Постановка задачі та варіант завдання

Номер варіанту завдань: 1
Завдання поставлені перед виконанням комп'ютерного парктикуму:

-1. Ознайомитись з порядком виконання комп'ютерного практикуму та відповідними вимогами до виконання роботи.
0. Уважно прочитати необхідні теоретичні відомості до комп'ютерного практикуму.
1. Створити новий репозиторій в системі контролю версій Git (бажано використовувати вебсервіс GitHub).
2. На великому тексті українською мовою (>1MB), необхідно розрахувати частоти літер і біграм, а також ентропію та індекс відповідності.
3. Отримати N текстів X українською мовою для довжин $L = 10, 100, 1000$ та 10000, для кожного з яких згенерувати спотворені тексти Y. Число N визначається відповідно до такої таблиці.

L	N
10 100 1000	10000
10000	1000

Спотворення тексту виконується такими способами:
(а) шляхом застосування шифру Віженера з випадковим ключем довжини r = 1, 5, 10.
(б) шляхом застосування шифру афінної та афінної біграмної підстановки з випадковими ключами.

А також тексти для аналізу формуються такими способами:
(а) y_i — рівномірно розподілена послідовність символів з $(Z_m)^l$
(б) y_i обчислюється відповідно до такого співвідношення:

$$y_i = (s_{i-1} + s_{i-2}) \mod m^l,$$

де $s_0, s_1 \in_R (Z_m)^l$.

4. Реалізувати критерії (відповідно до варіанту + структурний) і перевірити їх роботу на згенерованих N текстах для кожної довжини L. Розрахувати ймовірності похибок першого і другого роду. Номер варіанту Критерії

Номер варіанту	Критерії
Парний	1.0-1.3, 3.0, 5.1
Непарний	2.0-2.3, 4.0, 5.0

Усі вищезгадані критерії (та інші формули), які використовували значення l, мають приймати значення $l = 1$ та $l = 2$, тобто реалізувати символьний та біграмний критерії.

5. Згенерувати випадковий текст довжини $L = 10000$, який точно не є зв'язним текстом українською мовою (наприклад, текст, який складається з величезної кількості літер а: "ааааааааа..."). Застосувати один з варіантів спотворення (на вибір) до цього тексту, після чого застосувати один з реалізованих критеріїв (на вибір). Порівняти результати застосування критерію до різних текстів.
6. Оформити звіт до комп'ютерного практикуму.

3 Хід роботи

1. Ознайомлення з методичними вказівками та вимогами щодо виконання комп'ютерного практикуму
2. Ознайомлення, вивчення та систематизація необхідного теоретичного матеріалу для виконання комп'ютерного практикуму
3. Створення та налаштування репозиторію
4. Безпосередня реалізація поставлених задач комп'ютерного практикуму за допомогою програмування
5. Аналіз та систематизація отриманих результатів
6. Підготовка звіту з виконання комп'ютерного практикуму з детальним описом отриманих результатів

4 Варіант №3

4.1 Опис множин заборонених/частих символів, які було отримано при виконанні завдання

Опишемо загальну характеристику, вибірку з п'яти найчасчтіших символів та символів, які зустрічаються дуже рідко:

- **Найчастіші символи: о, а, н, в, и**
- **Найрідкіші символи: щ, ц, є, ї, ф**

4.2 Таблиці результатів тестів

- $\alpha = P(H_1|H_0)$ - ймовірність помилки 1-го роду (англ.*false positive*), ймовірність назвати відкритий текст випадковою послідовністю;
- $\beta = P(H_0|H_1)$ - ймовірність помилки 2-го роду (англ.*false negative*), ймовірність прийняти випадкову послідовністю за відкритий текст.

Шифрування за Віженером (r = 1)					
		Монограма		Біграма	
L	Номер критерію	$\alpha(l=1)$	$\beta(l=1)$	$\alpha(l=2)$	$\beta(l=2)$
10	2.0 60000/9000	43,6%	5,51%	85,6%	0,63%
	2.1 40000/9000,$k_{(f_1)}$=11, $k_{(f_2)}$=400	87,68%	47,22%	96,19%	31,5%
	2.2 40000/9000,$k_{(x_1)}$=2, $k_{(x_2)}$=1	65,21%	33%	87,67%	1,89%
	2.3 40000/9000,$K_{(f_1)}$=11, $K_{(f_2)}$=7	38,76%	17,26%	87,67%	3,24%
	4.0 0,02/0,04,$k_{(I_1)}$=5,2, $k_{(I_2)}$=0,9	46,71%	53,32%	100%	0%
	5.0 4000/50,$k_{(empt_1)}$=5, $k_{(empt_2)}$=40	2,03%	73,2%	0%	22,8%
100	2.0 15000/7000	25,8%	0%	15,33%	4,33%
	2.1 15000/7000,$k_{(f_1)}$=3, $k_{(f_2)}$=520	25,8%	0%	15,33%	4,4%
	2.2 15000/7000,$k_{(x_1)}$=6, $k_{(x_2)}$=2	25,8%	0%	18,41%	10,26%
	2.3 15000/7000,$K_{(f_1)}$=320, $K_{(f_2)}$=40	20,88%	0%	23,41%	4,22%
	4.0 0,02/0,004,$k_{(I_1)}$=4,87, $k_{(I_2)}$=0,92	94,66%	15,4%	100%	0%
	5.0 4000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=41,25	20,05%	55,6%	0%	0%
1000	2.0 20000/7000	0%	93,98%	0%	15,3%
	2.1 2000/6000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	2,5%	78,14%	0%	39,54%
	2.2 15000/7000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	25,5%	0%	15,98%	0%
	2.3 15000/7000,$K_{(f_1)}$=857, $K_{(f_2)}$=35	2,5%	0%	1%	28,16%
	4.0 0,00002/0,0003,$k_{(I_1)}$=0,005, $k_{(I_2)}$=0,015	66,6%	24,5%	0%	1%
	5.0 7000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=0,54	67,45%	0%	0%	0%
10000	2.0 20000/5000	100%	0%	86,43%	22%
	2.1 2000/5000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	100%	0%	0%	15,33%
	2.2 15000/5000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	45,24%	56,32%	0%	20,33%
	2.3 15000/5000,$K_{(f_1)}$=800, $K_{(f_2)}$=35	37,12%	0%	0%	12,33%
	4.0 0,005/0,002,$k_{(I_1)}$=0,015, $k_{(I_2)}$=0,015	45,3%	55,7%	25,5%	15,3%
	5.0 80000/60,$k_{(empt_1)}$=1, $k_{(empt_2)}$=200	100%	0%	0%	0%

Шифрування за Віженером (r = 5)					
		Монограма		Біграма	
L	Номер критерію	**FP**$(l=1)$	**FN**$(l=1)$	**FP**$(l=2)$	**FN**$(l=2)$
10	2.0 60000/9000	43,6%	15,51%	85,6%	0,63%
	2.1 40000/9000,$k_{(f_1)}$=11, $k_{(f_2)}$=400	87,68%	19,1%	88,9%	1%
	2.2 40000/9000,$k_{(x_1)}$=2, $k_{(x_2)}$=1	65,21%	29,65%	88,89%	1,5%
	2.3 40000/9000,$K_{(f_1)}$=11, $K_{(f_2)}$=7	38,76%	26,26%	78,76%	2,55%
	4.0 0,02/0,04,$k_{(I_1)}$=5,2, $k_{(I_2)}$=0,9	46,71%	31,78%	100%	0%
	5.0 4000/50,$k_{(empt_1)}$=5, $k_{(empt_2)}$=40	2,03%	75,37%	4,2%	15,57%
100	2.0 15000/7000	25,8%	22,4%	-%	-%
	2.1 15000/7000,$k_{(f_1)}$=3, $k_{(f_2)}$=520	25,8%	22,46%	22,43%	12,3%
	2.2 15000/7000,$k_{(x_1)}$=6, $k_{(x_2)}$=2	25,8%	31,28%	22,43%	15,98%
	2.3 15000/7000,$K_{(f_1)}$=320, $K_{(f_2)}$=40	20,88%	0%	12,43%	5,37%
	4.0 0,02/0,004,$k_{(I_1)}$=4,87, $k_{(I_2)}$=0,92	94,66%	10,55%	100%	0%
	5.0 4000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=41,25	20,05%	2,07%	21,5%	0%
1000	2.0 20000/7000	0%	100%	0%	15,39%
	2.1 2000/6000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	2,45%	100%	0%	22,12%
	2.2 15000/7000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	28,88%	0%	12,03%	0%
	2.3 15000/7000,$K_{(f_1)}$=8570, $K_{(f_2)}$=35	8,32%	0%	0%	22,88%
	4.0 0,00002/0,0003,$k_{(I_1)}$=0,005, $k_{(I_2)}$=0,015	3,17%	0%	0%	1,73%
	5.0 7000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=0,54	75%	0%	3,16%	0%
10000	2.0 20000/5000	100%	0%	100%	0%
	2.1 2000/5000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	100%	0%	0%	25,55%
	2.2 15000/5000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	100%	0%	0%	0%
	2.3 15000/5000,$K_{(f_1)}$=800, $K_{(f_2)}$=35	34,52%	0%	11,44%	24,76%
	4.0 0,005/0,002,$k_{(I_1)}$=0,015, $k_{(I_2)}$=0,015	100%	0%	25,22%	0%
	5.0 10000/60,$k_{(empt_1)}$=1, $k_{(empt_2)}$=200	100%	0%	0%	0%

Шифрування за Віженером (r = 10)					
		Монограма		Біграма	
L	Номер критерію	**FP**$(l=1)$	**FN**$(l=1)$	**FP**$(l=2)$	**FN**$(l=2)$
10	2.0 60000/9000	43,6%	29,91%	85,31%	0%
	2.1 40000/9000,$k_{(f_1)}$=11, $k_{(f_2)}$=400	87,68%	37,22%	76,14%	12,56%
	2.2 40000/9000,$k_{(x_1)}$=2, $k_{(x_2)}$=1	65,21%	32,04%	78,42%	0%
	2.3 40000/9000,$K_{(f_1)}$=11, $K_{(f_2)}$=7	38,76%	28,26%	78,89%	1,64%
	4.0 0,02/0,04,$k_{(I_1)}$=5,2, $k_{(I_2)}$=0,9	46,71%	38,13%	0%	100%
	5.0 4000/50,$k_{(empt_1)}$=5, $k_{(empt_2)}$=40	2,03%	73,78%	0%	21,22%
100	2.0 15000/7000	25,8%	28,05%	12,56%	12,71%
	2.1 15000/7000,$k_{(f_1)}$=3, $k_{(f_2)}$=520	25,8%	35,29%	0%	10,18%
	2.2 15000/7000,$k_{(x_1)}$=6, $k_{(x_2)}$=2	25,8%	35,47%	12,56%	12,89%
	2.3 15000/7000,$K_{(f_1)}$=320, $K_{(f_2)}$=40	20,88%	0%	15,67%	0%
	4.0 0,02/0,004,$k_{(I_1)}$=4,87, $k_{(I_2)}$=0,92	4,66%	5,99%	15,7%	100%
	5.0 4000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=41,25	20,05%	2,54%	0%	0%
1000	2.0 20000/7000	100%	0%	40,5%	0%
	2.1 2000/6000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	0%	100%	0%	12,54%
	2.2 15000/7000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	35,83%	15,32%	0%	6,5%
	2.3 15000/7000,$K_{(f_1)}$=857, $K_{(f_2)}$=35	10,66%	0%	0%	15,58%
	4.0 0,00002/0,0003,$k_{(I_1)}$=0,005, $k_{(I_2)}$=0,015	0%	0%	0%	0%
	5.0 7000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=0,54	89,8%	0%	3,17%	0%
10000	2.0 20000/5000	100%	0%	15,72%	55,55%
	2.1 2000/5000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	100%	0%	100%	0%
	2.2 15000/5000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	5,99%	5,68%	0%	0%
	2.3 15000/5000,$K_{(f_1)}$=800, $K_{(f_2)}$=35	0%	0%	0%	55,55%
	4.0 0,005/0,002,$k_{(I_1)}$=0,015, $k_{(I_2)}$=0,015	0%	0%	0%	0%
	5.0 30000/60,$k_{(empt_1)}$=1, $k_{(empt_2)}$=200	100%	0%	11%	0%

Шифрування за допомогою афінної підстановки з ключами a=5, b=7					
		Монограма		Біграма	
L	Номер критерію	**FP**$(l=1)$	**FN**$(l=1)$	**FP**$(l=2)$	**FN**$(l=2)$
10	2.0 60000/9000	43,5%	33,33%	88,43%	1,1%
	2.1 40000/9000,$k_{(f_1)}$=11, $k_{(f_2)}$=400	26,9%	15,73%	26,93%	5,52%
	2.2 40000/9000,$k_{(x_1)}$=2, $k_{(x_2)}$=1	43,5%	33,33%	88,43%	1,4%
	2.3 40000/9000,$K_{(f_1)}$=11, $K_{(f_2)}$=7	20,02%	10%	31,4%	5,76%
	4.0 0,02/0,04,$k_{(I_1)}$=5,2, $k_{(I_2)}$=0,9	58,8%	42,4%	31,78%	57,9%
	5.0 4000/50,$k_{(empt_1)}$=5, $k_{(empt_2)}$=40	3,21%	95,6%	7,22%	6,31%
100	2.0 15000/7000	1%	7,8%	25,8%	12,566%
	2.1 15000/7000,$k_{(f_1)}$=3, $k_{(f_2)}$=520	0%	100%	0%	22,2%
	2.2 15000/7000,$k_{(x_1)}$=6, $k_{(x_2)}$=2	1%	7,8%	25,8%	12,56%
	2.3 15000/7000,$K_{(f_1)}$=320, $K_{(f_2)}$=40	0%	1,77%	1,37%	25,44%
	4.0 0,02/0,004,$k_{(I_1)}$=4,87, $k_{(I_2)}$=0,92	50,05%	50,44%	56,53%	77,52%
	5.0 4000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=41,25	20,06%	50,44%	1%	0%
1000	2.0 20000/7000	0%	88,84%	0%	21,32%
	2.1 2000/6000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	0%	88,84%	0%	3,65%
	2.2 15000/7000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	0%	1%	0%	21,32%
	2.3 15000/7000,$K_{(f_1)}$=857, $K_{(f_2)}$=35	0%	1%	0%	0%
	4.0 0,00002/0,0003,$k_{(I_1)}$=0,005, $k_{(I_2)}$=0,015	54,44%	45,56%	27,5%	10,76%
	5.0 7000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=0,54	98,89%	1,38%	2,67%	0%
10000	2.0 20000/5000	100%	0%	79,87%	10,65%
	2.1 2000/5000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	10,65%	0%	0%	11,2%
	2.2 15000/5000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	0%	10,65%	0%	0%
	2.3 15000/5000,$K_{(f_1)}$=800, $K_{(f_2)}$=35	0%	10,55%	0%	10%
	4.0 0,005/0,002,$k_{(I_1)}$=0,015, $k_{(I_2)}$=0,015	1,5%	9,54%	0%	0%
	5.0 15000/60,$k_{(empt_1)}$=1, $k_{(empt_2)}$=200	10,65%	0%	0%	0%

Шифрування за рівномірно розподіленою послідовністю					
		Монограма		Біграма	
L	Номер критерію	**FP**$(l=1)$	**FN**$(l=1)$	**FP**$(l=2)$	**FN**$(l=2)$
10	2.0 60000/9000	41,56%	0%	85,23%	0%
	2.1 40000/9000,$k_{(f_1)}$=11, $k_{(f_2)}$=400	2,3%	0%	65,32%	2,55%
	2.2 40000/9000,$k_{(x_1)}$=2, $k_{(x_2)}$=1	41,56%	0%	85,23%	3,53%
	2.3 40000/9000,$K_{(f_1)}$=11, $K_{(f_2)}$=7	43,8%	0%	30,84%	22,24%
	4.0 0,02/0,04,$k_{(I_1)}$=5,2, $k_{(I_2)}$=0,9	55,38%	44,76%	12,45%	21,98%
	5.0 4000/50,$k_{(empt_1)}$=5, $k_{(empt_2)}$=40	2,47%	100%	10,2%	5,36%
100	2.0 15000/7000	0%	0%	12,44%	31,69%
	2.1 15000/7000,$k_{(f_1)}$=3, $k_{(f_2)}$=520	0%	0%	5,87%	0%
	2.2 15000/7000,$k_{(x_1)}$=6, $k_{(x_2)}$=2	0%	0%	23,05%	43,61%
	2.3 15000/7000,$K_{(f_1)}$=320, $K_{(f_2)}$=40	4%	0%	0%	1%
	4.0 0,02/0,004,$k_{(I_1)}$=4,87, $k_{(I_2)}$=0,92	54,87%	58,77%	88,53%	51,82%
	5.0 4000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=41,25	1,45%	100%	51,25%	0%
1000	2.0 20000/7000	0%	0%	1%	0%
	2.1 2000/6000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	0%	0%	1%	0%
	2.2 15000/7000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	0%	0%	1%	0%
	2.3 15000/7000,$K_{(f_1)}$=857, $K_{(f_2)}$=35	0%	0%	0%	5,77%
	4.0 0,00002/0,0003,$k_{(I_1)}$=0,005, $k_{(I_2)}$=0,015	53,56%	76,45%	78,45%	0%
	5.0 7000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=0,54	89,93%	100%	11,33%	0%
10000	2.0 20000/5000	100%	0%	0%	10%
	2.1 2000/5000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	100%	0%	10%	0%
	2.2 15000/5000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	5,3%	0%	0%	10%
	2.3 15000/5000,$K_{(f_1)}$=800, $K_{(f_2)}$=35	5,3%	0%	10%	0%
	4.0 0,005/0,002,$k_{(I_1)}$=0,015, $k_{(I_2)}$=0,015	1,4%	9%	0%	0%
	5.0 25000/60,$k_{(empt_1)}$=1, $k_{(empt_2)}$=200	100%	0%	0%	0%

Шифрування за псевдовипадковою послідовністю					
		Монограма		Біграма	
L	Номер критерію	**FP**$(l=1)$	**FN**$(l=1)$	**FP**$(l=2)$	**FN**$(l=2)$
10	2.0 60000/9000	54,32%	17,32%	87,59%	1%
	2.1 40000/9000,$k_{(f_1)}$=11, $k_{(f_2)}$=400	44,6%	27,89%	100%	0%
	2.2 40000/9000,$k_{(x_1)}$=2, $k_{(x_2)}$=1	44,6%	27,89%	88,56%	0%
	2.3 40000/9000,$K_{(f_1)}$=11, $K_{(f_2)}$=7	44,6%	27,89%	88,56%	0%
	4.0 0,02/0,04,$k_{(I_1)}$=5,2, $k_{(I_2)}$=0,9	20,66%	65,93%	21,54%	1%
	5.0 4000/50,$k_{(empt_1)}$=5, $k_{(empt_2)}$=40	0%	100%	6,22%	4,66%
100	2.0 15000/7000	1,33%	5,33%	22,67%	5,53%
	2.1 15000/7000,$k_{(f_1)}$=3, $k_{(f_2)}$=520	1,33%	5,89%	13,23%	1,5%
	2.2 15000/7000,$k_{(x_1)}$=6, $k_{(x_2)}$=2	1,33%	5%	22,5%	0%
	2.3 15000/7000,$K_{(f_1)}$=320, $K_{(f_2)}$=40	0%	15,65%	15,5%	0%
	4.0 0,02/0,004,$k_{(I_1)}$=4,87, $k_{(I_2)}$=0,92	31%	47,31%	1%	18,9%
	5.0 4000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=41,25	1%	99%	1%	0%
1000	2.0 20000/7000	0%	5,64%	0%	2,44%
	2.1 2000/6000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	0%	5,64%	0%	0%
	2.2 15000/7000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	0%	5,5%	0%	0%
	2.3 15000/7000,$K_{(f_1)}$=857, $K_{(f_2)}$=35	0%	0%	0%	0%
	4.0 0,00002/0,0003,$k_{(I_1)}$=0,005, $k_{(I_2)}$=0,015	1%	0%	0%	0%
	5.0 7000/50,$k_{(empt_1)}$=3, $k_{(empt_2)}$=0,54	78,5%	91,8%	2,53%	0%
10000	2.0 20000/5000	100%	5,64%	88,42%	0%
	2.1 2000/5000,$k_{(f_1)}$=7, $k_{(f_2)}$=250	100%	5,64%	0%	5,3%
	2.2 15000/5000,$k_{(x_1)}$=20, $k_{(x_2)}$=25	0%	5,63%	0%	0%
	2.3 15000/5000,$K_{(f_1)}$=800, $K_{(f_2)}$=35	0%	15,87%	0%	0%
	4.0 0,005/0,002,$k_{(I_1)}$=0,015, $k_{(I_2)}$=0,015	46,24%	100%	0%	9,4%
	5.0 100000/60,$k_{(empt_1)}$=1, $k_{(empt_2)}$=200	100%	10%	0%	0%

4.3 Структурний критерій

Алгоритм стиснення даних, реалізований у бібліотеці zlib Python, використовує модифікацію алгоритму DEFLATE, який є основою для багатьох популярних архіваторів. Була написана функція `calculate_compression_ratio`, яка обраховує коефіцієнт стиснення заданого тексту. Функція приймала на вхід текст та довжину тексту. Логіка стиснення в тому, що алгоритм стиснення zlib перетворює текст в байти з використанням кодировки UTF-8. Це необхідно, тому що zlib працює з байтами, а не з рядками тексту . Для побудови структурного критерію використовувася функція `structure_criteria`. Для кожного рядка порівнюються коефіцієнты стиснення випадкової послідовності і відповідного реального рядка. Якщо різниця між ними більше за порогове значення limit рахуємо, що текст є структурованим, тобто приймає гіпотезу H_1, інакше приймаємо H_0.

4.4 Опис труднощів

Основними труднощами особисто для мене стали великий об'єм роботи і підбір порогових значень для критеріїв. Також виникали труднощі із написанням функцій для обрахунку критеріїв, але їх було подолано.

4.5 Висновки

В ході цій роботі було програмно реалізовано алгоритми для шфирування текстів і генерації випадкових послідовностей. А також було реалізовано критерії перевірки гіпотез, які визначають, чи є вхідна послідовність осмимленим тектом чи випадковою послідовністю. Окрім цього були реалізовані функції стиснення та струкутурний критерій, який використовує алгоритм стиснення DEFLATE.Результатом роботи є таблиці з оцінками помилок першого та другого роду.Проаналізувавши результати, виявили, що критерії1 2.0 та 2.1 мають найбільш ефективні реузультати для невеликих довжин послідовностей L=10 та L=100. Це цілком очевидно, так як при великих довжинах тексту більшість частотних біграм, ймовірно, з'являться в тексті, а ймовірність появи монограм буде ще вищою. Критерій 4.0 навпаки працює краще на довжинах L=1000 та L=10000. Структурний критерій, застосований до відкритого тексту довжиною L=1000 символів і відпоідного шифротексту за Віженером з ключем 5, показав, що помилки першого та другого роду будуть 0, так як критерій у всіх випадках визначив відкритий текст за відкритий. Однак для текстів довжиною L=10 та L=100 критерій видає гірші результати.