МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО»

Фізико-технічний інститут

Криптографія

Комп’ютерний практикум №1

Виконали:

Студенти групи ФБ-05

Сапожник М.В.

Карась Б.І.

Київ - 2022

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

**Постановка задачі:**

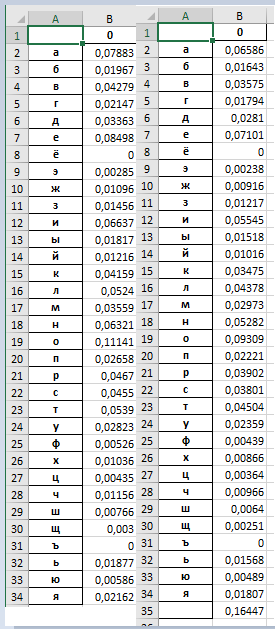
Написати програми для підрахунку частот букв і частот біграм в тексті, також для підрахунку H1 та H2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення H1 та H2 на довільно обраному тексті російською мовою достатньої довжини, де імовірності замінити відповідними частотами. Одержати значення H1 та H2 на тому ж тексті, в якому вилучено всі пробіли.

**Хід роботи:**

Спершу для знайденого тексту російською мовою провели обробку(перевели усі символи в нижній регістр, видалили повторні пробіли та усі символи крім зазначених),отримали два очищених файли(один з пробілами інший без) та дві відповідні змінні. Також реалізували функції показування частоти біграм та частоти перехресних біграм.

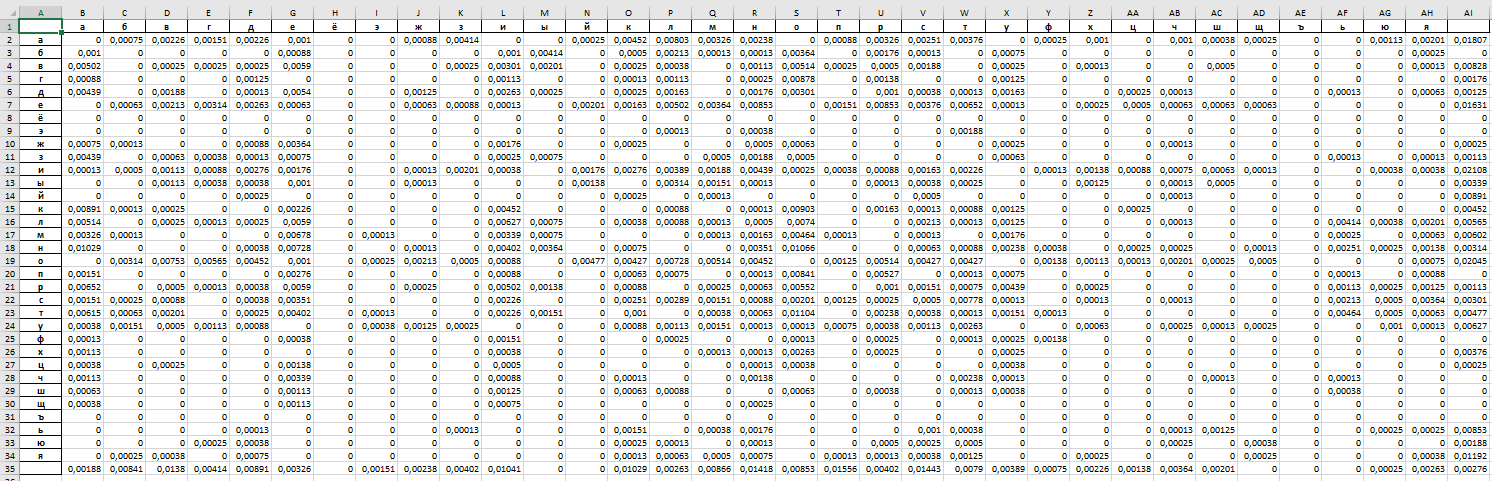
Далі написали функції для підрахунку ентропії та надлишковості, використовуючи функції наведені у методичці до комп’ютерного практикуму

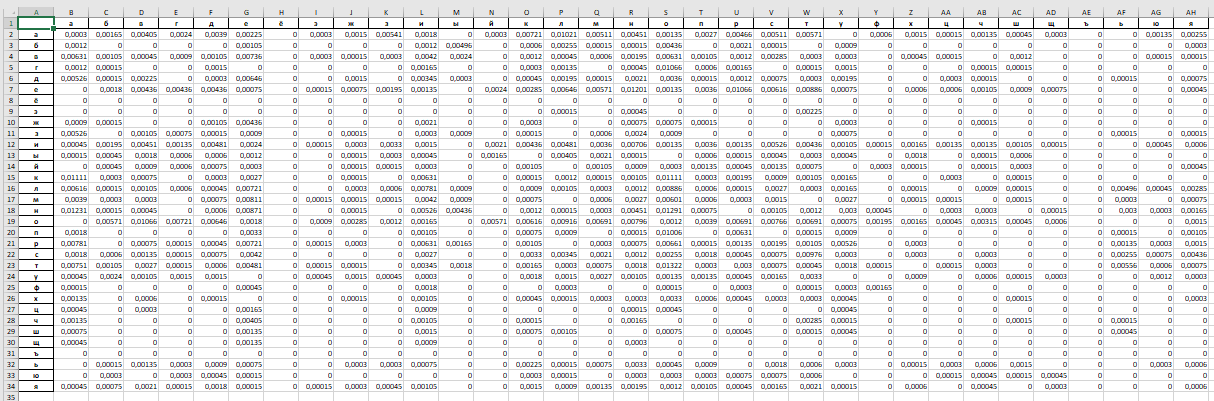
Після цього реалізували функцію, яка заносить дані у .xlsx табличку. На цьому етапі виникли невеличкі проблеми: коли ми реалізували вивід у .xlsx табличку, треба було завантажити певні модулі для того, щоб сам код працював без помилок, та виводив данні у найменовані для них таблиці( код виконувався і виводив у командній строчці результати, проте після них виводилися помилки)

****

Перша табличка демонструє частоти для букв у тексті без пробілів, друга частоти для букв у тексті з пробілами.

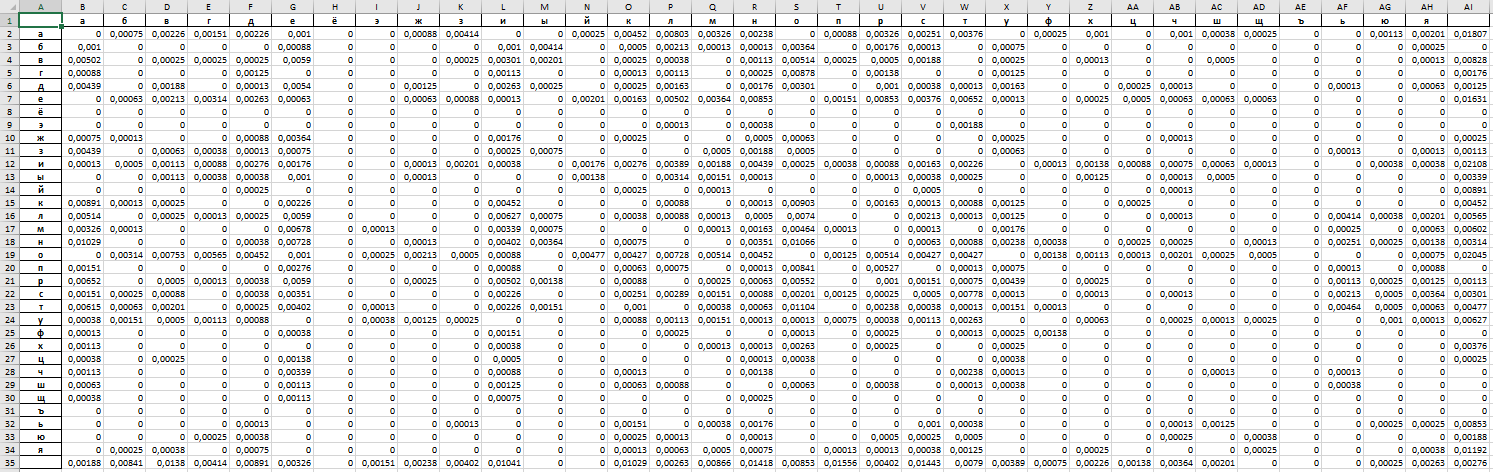
**Таблиці частот біграм:**

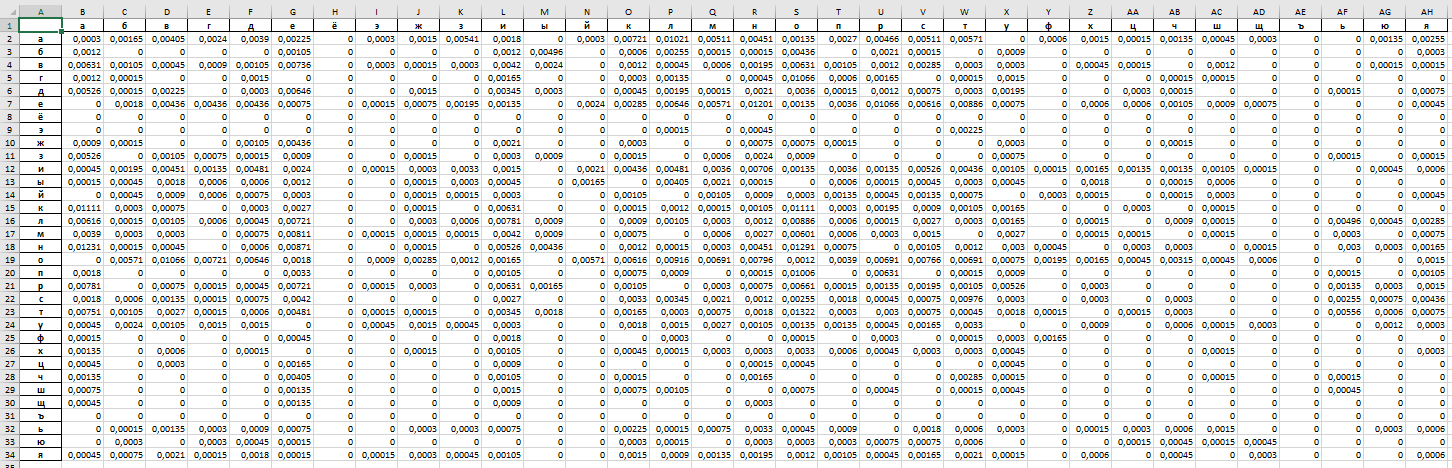
****

****

На першій таблиці біграми з пробілами, на другій таблиці, біграми без пробілів

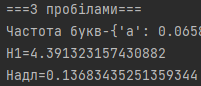
**Таблиця частот перехресних біграм:**

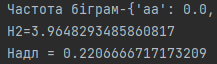


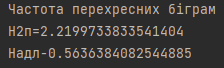


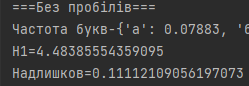
На першій таблиці перехресні біграми з пробілами, на другій таблиці перехресні біграми без пробілів

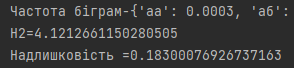
**Одержані значення H та надлишку:**

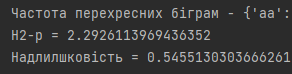




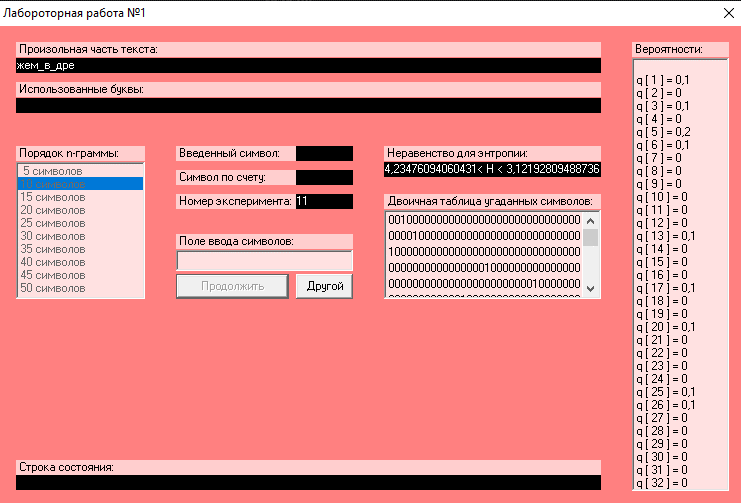


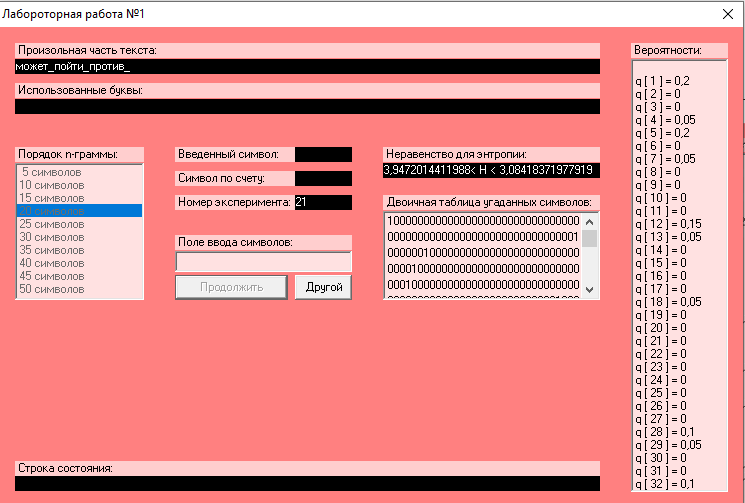


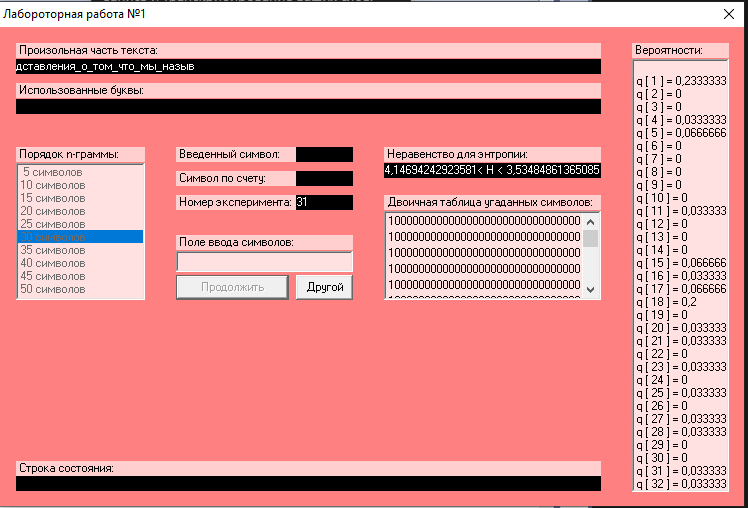




Оцінки для H(10), H(20), H(30):







Висновки:

Виконавши цей комп’ютерний практикум ми набули практичних навичок щодо оцінки ентропії на символ джерела, порівняли різні моделі джерела відкритого тексту для наближеного визначення ентропії.

Написали програму для підрахунку усіх необхідних для роботи значень, використавши для збереження отриманих результатів модуль, який надає змогу запису даних у .xlsx файл.

Використавши отримані значення ентропії, підрахували надлишковість. За допомогою CoolPinkProgram оцінили значення для H(10), H(20), H(30).