Q1) Sample a chain of 5000 integers with 6 numbers (1,2,3,4,5,6).

Q2) For any particular integer (say 1), consider only the next number in the sequence you obtained in Question 2 (call this as the sample space for integer 1). You have to calculate the sample space for every integer in the sequence. For example, suppose you have the following sequence of three numbers.

1,"2",3,3,2,3,1,"2",3,2,1

The sample space for each of the numbers is as follows:

Number (x) Sample space for x Sample space

1 2,2 (Look for numbers between " ") S1

2 3,3,3 S2

3 3,2,1,2 S3

Q3) For the sample spaces (S1-S6) obtained in the previous question, calculate the probability of each element in each of sample spaces respectively.

Q4) Calculate the frequency of n-grams for the old file n=2,3,4 (Lab.csv). In this question, you have to consider n words at a time.