

Лабораторная работа № 1 “Hello World of Python”.

Теоретическая часть (Теория вероятностей):

1. (1 балл) Рассчитайте математическое ожидание случайной величины X , принимающей значения $\{-1, 0, 1\}$ с вероятностями, соответственно, $\{0.3, 0.2, 0.5\}$.
2. (1 балл) Известно, что $E(X) = 4$, $E(Y) = 1$, $\text{Var}(X) = 4$, $\text{Var}(Y) = 1$, $\text{Cov}(X, Y) = 1$. Найдите $E(X+2Y)$.
3. (1 балл) Известно, что $E(X) = 1$, $E(Y) = 3$, $\text{Var}(X) = 4$, $\text{Var}(Y) = 1$, $\text{Cov}(X, Y) = 1$. Найдите $\text{Var}(X+2Y)$.
4. (3 балла) Игральную кость бросают до тех пор, пока не выпадет 1. Найти вероятность того, что это случилось на втором бросании, если известно, что для этого потребовалось чётное число бросаний.

Практическая часть:

Все задания желательно выполнить без использования циклов (за исключением задания 6).

1. (Вектор) (5 баллов) Создать векторы $v1$ и $v2$:

Вектор $v1$ состоит из последовательных членов арифметической прогрессии. Первый член прогрессии равен вашему номеру в списке группы, последний – числу 6.5, а количество элементов последовательности – длине вашей фамилии.

Вектор $v2$ состоит из последовательных членов геометрической прогрессии. Пусть ваш номер в списке i . Первый член прогрессии равен $|13-i|+2$, последний – числу 100, а количество элементов последовательности – длине вашего имени.

Вывести вектор $v3$ длины 3, состоящий из случайно выбранных элементов векторов $v1$ и $v2$.

2. (Вектор) (5 баллов) Задана некоторая строка текста, состоящая только из строчных символов русского языка. Например $s <-$ “приветмир”. Алфавит можно задать в явном виде ($alf <-$ “абвг...”).

1 вариант. Найти сумму порядковых номеров в алфавите символов строки s . Нумерация букв алфавита начинается с 1. Например, для $s =$ “абвг” ответ должен быть 10.

2 вариант. Найти произведение тех порядковых номеров в алфавите символов строки s , номера которых нечетные. Нумерация букв алфавита начинается с 1. Например, для $s =$ “абвг” ответ должен быть 3.

3 вариант. Зашифровать текст s с помощью шифра Цезаря, т.е. такого шифра, при котором i -ая буква алфавита переходит в $((i+3) \% N)$ -ую букву алфавита, где N – количество букв в алфавите. Например, для $s =$ “абвг” ответ должен быть “гдеё”.

4 вариант. 1-ую букву строки s оставить без изменений, 2-ую – заменить на следующую в алфавите, 3-ую – заменить на стоящую через одну в алфавите и т.д. Например, $s =$ “абвг” ответ должен быть “авдё”.

5 вариант. Преобразовать строку s таким образом, чтобы i -ая буква алфавита перешла в $(N - i)$ -ую. Например, $s =$ “абвг” ответ должен быть “яюэь”.

Указание. Могут пригодиться следующие функции: `match`, `substring`, `strsplit`.

3. (Матрица) (5 баллов) Создать квадратную матрицу m размером 5×5 . Заполнить ее элементами вектора $v1$ построчно, если первый элемент $v3$ больше 10, иначе по столбцам. Если длины $v1$ не хватает, то пусть элементы этого вектора повторяются. Присвоить столбцам имена следующим образом: если N – это ваш номер по списку, то n равно остатку от деления $(N-1)$ на 12 плюс 1, тогда первый столбец имеет имя n -го месяца на английском языке, второй – $(n+1)$ -го и т.д. Строки присвоить такие же имена как и столбцам. Матрица $m1$ получается из матрицы m путем вычеркивания строк и столбцов, чьи имена начинаются на буквы от “A” до “F”.

Для матрицы $m1$ найти: определитель, собственные вектора и значения, вектор диагональных элементов, $v1^2$ и матрицу $v2$, у которой $v2[i][j] = (v1[i][j])^2$

4. (Таблица) (5 баллов) Создать произвольную таблицу данных, в которой будут присутствовать данные следующих типов: числовые, текстовые, условные, факторы. Вывести все такие текстовые значения для заданного фактора, для которых числовые значения больше заданного значения.

5. (Таблица) (5 баллов) Из файла загрузить таблицу, состоящую из трех переменных (x_1, x_2, x_3) (переменная = столбец) и n наблюдений (строк).

Вычислить

1 вариант $\sum_{i=1}^n i x_{1,i}^2 x_{2,i \% n+1}^4 x_{2,(i+1) \% n+1}^3$;

2 вариант $\sum_{i=1}^n ((i-1) \% 5 + 1) x_{(i-1) \% 3+1,i}^i$;

3 вариант $\sum_{i=1}^n \left(\frac{x_{1,i}}{x_{2,i}} - x_{3,i} \right)^2 (x_{1,1} + x_{(i-1) \% 3+1,1})$;

4 вариант $\sum_{i=1}^n (I\{x_{1,i} < x_{2,i} \text{ и } x_{1,i} < x_{3,i}\} (x_{3,i} - x_{1,i})(x_{2,i} - x_{1,i}) + I\{x_{2,i} > x_{3,i}\} x_{1,i}^2);$

5 вариант $\sum_{j=1}^3 \sum_{i=1}^n I\{x_{j,i} \leq x_{1,i} \text{ и } x_{j,i} \leq x_{2,i} \text{ и } x_{j,i} \leq x_{3,i}\} x_{j \% 3 + 1, i}^{(i+j) \% 3 + 1};$

6 вариант $\sum_{j=1}^n (x_{1,i} + x_{2,i} + x_{3,i}^2) * (x_{3, n-i+1} + x_{2, n-i+1} + x_{1, n-i+1}^2);$

7 вариант $\sum_{i=1}^n (x_{1,i} + x_{2, n-i+1} + x_{3,i}^i).$

6. Задание по желанию:

Huffman coding (https://en.wikipedia.org/wiki/Huffman_coding)

6.1 (15 баллов)

По данной непустой строке s длины не более 10^4 , состоящей из строчных букв латинского алфавита, постройте оптимальный беспрефиксный код. В первой строке выведите количество различных букв k , встречающихся в строке, и размер получившейся закодированной строки. В следующих k строках запишите коды букв в формате "letter: code". В последней строке выведите закодированную строку.

Sample Input 1:

a

Sample Output 1:

1 1

a: 0

0

Sample Input 2:

abacabad

Sample Output 2:

4 14

a: 0

b: 10

c: 110

d: 111

01001100100111

6.2 (7 баллов)

Восстановите строку по её коду и беспрефиксному коду символов.

В первой строке входного файла заданы два целых числа k и l через пробел — количество различных букв, встречающихся в строке, и размер получившейся закодированной строки, соответственно. В следующих k строках записаны коды букв в формате "letter: code". Ни один код не является префиксом другого. Буквы могут быть перечислены в любом порядке. В качестве букв могут встречаться лишь строчные буквы латинского алфавита; каждая из этих букв встречается в строке хотя бы один раз. Наконец, в последней строке записана закодированная строка. Исходная строка и коды всех букв непусты. Заданный код таков, что закодированная строка имеет минимальный возможный размер.

В первой строке выходного файла выведите строку s . Она должна состоять из строчных букв латинского алфавита.

Sample Input 1:

1 1

a: 0

0

Sample Output 1:

a

Sample Input 2:

4 14

a: 0

b: 10

c: 110

d: 111

01001100100111

Sample Output 2:

abacabad