

Санкт-Петербургский Государственный Технический Университет  
(Технологический институт)

Кафедра системного анализа и информационных технологий

### **Лабораторная работа №5**

Выполнили:

Леякин А. А., Люцай В.Н.

Проверил

Мусаев А.А.

Санкт-Петербург,

2022

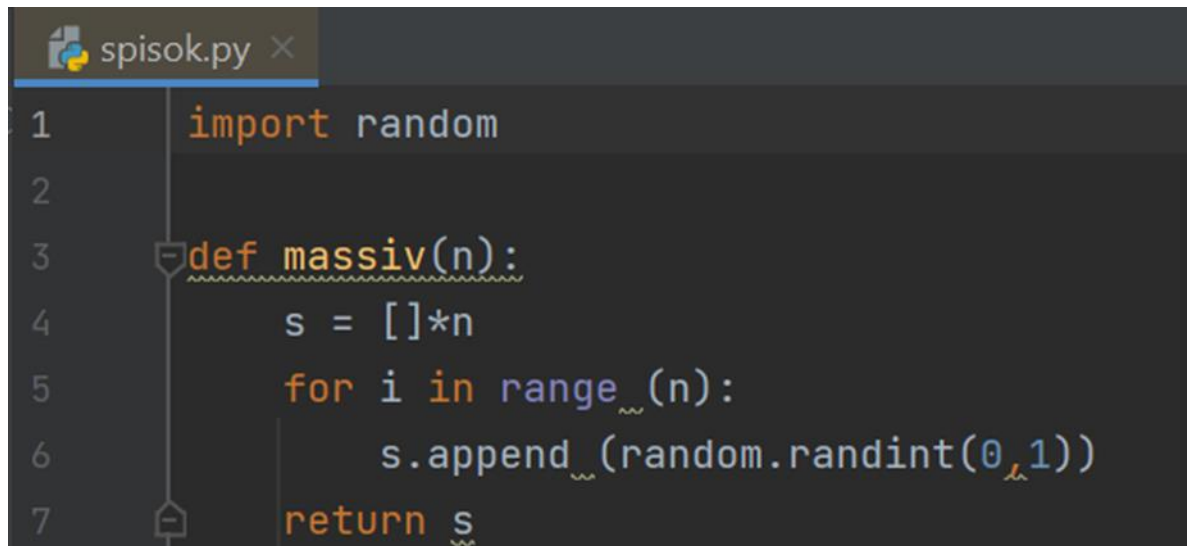
## Ход работы

### Задание 1

“Изучите модуль `random`. Создайте список из  $N$  случайных элементов, которые могут принимать значение 0 или 1. Для созданного списка определить процентное отношение 0, 1, а также повторяющихся подряд элементов (00, 11, 000, 111 и т.д.).”

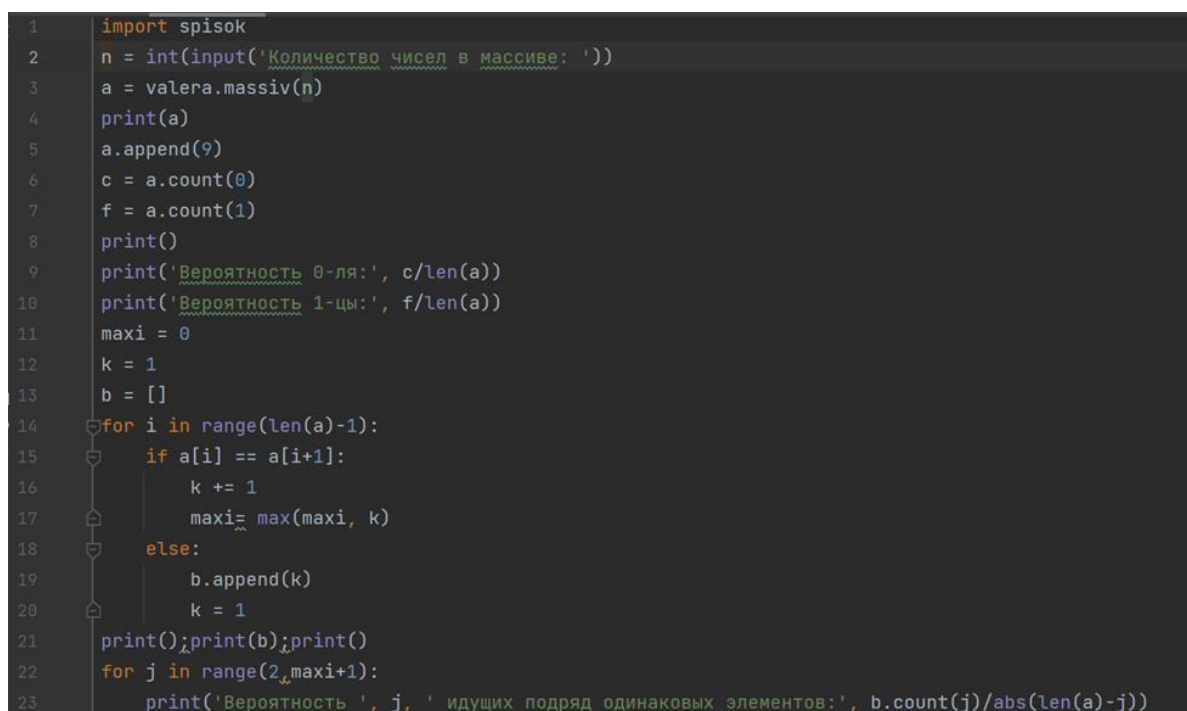
### Выполнение:

Мы написали программу, которая случайным образом (с помощью метода `random`) создает массив из нулей и единиц (длину массива задает пользователь). Программа определяет вероятность нуля или единицы, а также вероятность повторяющихся подряд элементов массива (рис. 1-2).



```
1 import random
2
3 def massiv(n):
4     s = []*n
5     for i in range(n):
6         s.append(random.randint(0,1))
7     return s
```

рис.1



```
1 import spisok
2 n = int(input('Количество чисел в массиве: '))
3 a = valera.massiv(n)
4 print(a)
5 a.append(9)
6 c = a.count(0)
7 f = a.count(1)
8 print()
9 print('Вероятность 0-ля:', c/len(a))
10 print('Вероятность 1-цы:', f/len(a))
11 maxi = 0
12 k = 1
13 b = []
14 for i in range(len(a)-1):
15     if a[i] == a[i+1]:
16         k += 1
17         maxi = max(maxi, k)
18     else:
19         b.append(k)
20         k = 1
21 print(); print(b); print()
22 for j in range(2, maxi+1):
23     print('Вероятность ', j, ' идущих подряд одинаковых элементов:', b.count(j)/abs(len(a)-j))
```

рис.2

## Задание 2

“Изучите модуль `matplotlib`. Для случайного набора данных, определить математическое ожидание, среднеквадратическое отклонение и построить линейную функцию, используя метод наименьших квадратов.”

### Выполнение:

Мы написали программу, которая определяет математическое ожидание и среднеквадратическое отклонение для введенного пользователем массива чисел (рис. 3-5), а также строит график в оси координат, где по оси ординат отложены значения чисел, а по оси абсцисс порядковые номера чисел в массиве. В этом графике синим цветом отмечены точки и линейная функция, построенная с помощью метода наименьших квадратов. Команда `.polyfit`, импортированная из библиотеки `matplotlib`, позволяет построить данную линейную зависимость.

```
1  import numpy as np
2  import math as ma
3  q = int(input('Введите количество чисел: '))
4  w = [0] * q
5  for i in range(q):
6      w[i] = int(input('Введите число: '))
7  print(w, '- заданный массив чисел')
8  y = list(set(w))
9  exp = 0
10 for j in range(len(y)):
```

рис.3

```
11     v = w.count(y[j])/len(w)
12     print('Вероятность выпадения', y[j], 'равна', v)
13     exp += (v * y[j])
14     print(exp, '- математическое ожидание')
15     hvalue = 0
16     for i in range(q):
17         h = (w[i] - exp)**2
18         hvalue += h
19     disp = hvalue/len(w)
20     print(disp, '- значение дисперсии')
21     otkl = ma.sqrt(disp)
22     print(otkl, '- среднеквадратическое отклонение')
23     import matplotlib.pyplot as plt
24     x1 = []
25     k = 1
26     for j in range(len(w)):
27         x1.append(k)
28         k = k + 1
29     y1 = w
30     c = np.polyfit(x1, y1, 1)
```

рис.4

```

31 ka = c[0]
32 kb = c[1]
33 x = np.linspace(0, len(w) + 1, 100)
34 y = ka * x + kb
35 plt.scatter(x1, y1)
36 plt.grid()
37 plt.plot(x, y)
38 plt.show()

```

рис.5

Пример работы программы (рис. 6-7):

```

Введите количество чисел: 5
Введите число: 2
Введите число: 4
Введите число: 2
Введите число: 5
Введите число: 8
[2, 4, 2, 5, 8] - заданный массив чисел
Вероятность выпадения 8 равна 0.2
Вероятность выпадения 2 равна 0.4
Вероятность выпадения 4 равна 0.2
Вероятность выпадения 5 равна 0.2
4.2 - математическое ожидание
4.9600000000000001 - значение дисперсии
2.227105745132009 - среднеквадратическое отклонение

```

рис.6

Figure 1

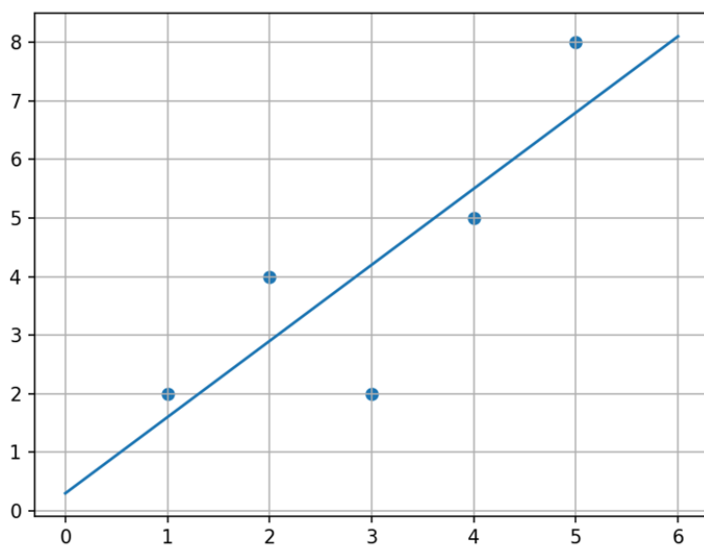


Figure 1

рис.7

### Задание 3

“Скачайте данные о котировках акций Visa и Mastercard за любой год с интервалом 1 день. Используя коэффициент корреляции Пирсона, оцените взаимосвязь этих двух компаний. Удалите случайные данные о котировках данных компаний и напишите функции для их восстановления с помощью винзорирования, линейной аппроксимации и корреляционного восстановления. Сделайте выводы о каждом методе восстановления данных.”

#### Выполнение:

Мы скачали данные о котировках акций Visa и Mastercard и расположили их в Excel (рис. 8):

	A	B	C	D	E	F
1	Дата	Цена	Visa	Дата	Цена	Mastercard
2	31.12.2021	216,71		31.12.2021	359,32	
3	30.12.2021	217,87		30.12.2021	360,99	
4	29.12.2021	218,17		29.12.2021	361,29	
5	28.12.2021	218,05		28.12.2021	360,78	
6	27.12.2021	217,63		27.12.2021	360,31	
7	23.12.2021	216,62		23.12.2021	360,58	
8	22.12.2021	217,96		22.12.2021	357,48	
9	21.12.2021	215,38		21.12.2021	350,74	
10	20.12.2021	208,74		20.12.2021	337,02	
11	17.12.2021	211,88		17.12.2021	349,31	
12	16.12.2021	214,37		16.12.2021	352,92	
13	15.12.2021	212,31		15.12.2021	346,35	
14	14.12.2021	209,98		14.12.2021	340,55	
15	13.12.2021	211,02		13.12.2021	345,31	
16	10.12.2021	213,4		10.12.2021	349,92	
17	09.12.2021	211,39		09.12.2021	344,52	
18	08.12.2021	208,99		08.12.2021	342,34	
19	07.12.2021	207,37		07.12.2021	339,51	
20	06.12.2021	202,68		06.12.2021	333,33	
21	03.12.2021	196,32		03.12.2021	322,11	
22	02.12.2021	198,29		02.12.2021	320,2	

рис.8

Взаимосвязь этих двух компаний можно оценить с помощью выборочного линейного коэффициента парной корреляции Пирсона. Для этого выберем свободную ячейку и, с помощью функции КОРРЕЛ, вычислим коэффициент корреляции (рис. 9):

Кэф. Коррел.
0,789667555

рис.9

Чтобы оценить взаимосвязь этих компаний воспользуемся шкалой Чеддока (рис. 10):

Диапазон значений $ r $	Линейная корреляционная зависимость $Y$ от $X$
0-0,1	практически отсутствует
0,1-0,3	слабая
0,3-0,5	умеренная
0,5-0,7	заметная
0,7-0,9	сильная
0,9-0,99	очень сильная
0,99-1	практически функциональная

рис.10

Сопоставив получившиеся значения коэффициента корреляции со шкалой Чеддока можно утверждать, что взаимосвязь компаний сильная.