



### Python Week 4: Mathematical Expectation and Correlation

In [ ]:

```
!python --version
```

Python 3.8.8

In [ ]:

```
import numpy as np
import scipy as sp
print(np.__name__, np.__version__)
print(sp.__name__, sp.__version__)
```

numpy 1.21.4

scipy 1.7.2

#### Exercise 1

A standard card deck consists of 52 cards, 13 cards for each of 4 suits: hearts, diamonds, spades, and clubs. Hearts and diamonds are red cards, while spades and clubs are black cards. In this exercise, we will simulate drawing 5 cards from the deck without replacement (i.e. after each draw, we do not replace the drawn card before drawing again), and count the number of clubs ( $C$ ), the number of spades ( $S$ ), the number of red cards ( $R$ ), and the number of black cards ( $B$ ) in the deck.

**a)** - Simulate 100,000 draws of 5 cards without replacement from the card deck, and record in vectors  $C$ ,  $S$ ,  $R$ , and  $B$  the numbers of clubs, spades, red cards, and black cards in the deck, respectively. You can use the **NumPy** package, to simulate 100,000 draws of 5 cards without replacement.

**b)** - Use the `mean()` function from the **NumPy** Package to compute the average number of clubs drawn, the average number of spades drawn, the average number of red cards drawn, and the average number of black cards drawn. Are these values what you expected?

**c)** - Use the `std( ,ddof=1)` function from the **NumPy** package to compute the standard deviation of the number of clubs drawn, the standard deviation of the number of spades drawn, the standard deviation of the number of red cards drawn, and the standard deviation of the number of black cards drawn.

**d)** - You have learnt that the correlation coefficient  $\rho$  is a measure of the association between two numerical variables, and that  $-1 \leq \rho \leq 1$ . Positive values of  $\rho$  indicate an increasing relationship between the two variables, while negative values of  $\rho$  indicate a decreasing relationship between the two variables. Consider the correlations between  $C$  and  $S$ , between  $C$  and  $R$ , between  $C$  and  $B$ , and between  $R$  and  $B$ . Which correlations would you expect to be positive? Which correlations would you expect to be negative?

**e)** - Use the `corrcoef()` function from the **NumPy** Package to compute the correlation coefficients between  $C$  and  $S$ , between  $C$  and  $R$ , between  $C$  and  $B$ , and between  $R$  and  $B$ . Do these correlations align with what you expected?