

**Thème :** Probabilité, statistique et informatique

**Titre :** Analyse de données issues de phénomènes naturels aléatoires.

**Auteur :** BRO FRÉDÉRIC

**Objectifs :**

- ▷ Programmer en PYTHON et utiliser notamment le module PANDAS pour étudier une étude statistique.
- ▷ Utiliser un intervalle de fluctuation pour tester une hypothèse.

On appelle BRUIT ATMOSPHERIQUE le bruit généré par les décharges électrostatiques qui sont produites par les orages électriques.

Le groupe de scientifiques de RANDOM.ORG collecte, analyse et traite les signaux sonores. Le principe consiste à :

- Utiliser une radio avec une fréquence où personne ne diffuse.
- Capturer le bruit atmosphérique à l'aide d'un récepteur dans un microphone.  
Ce signal sonore est ensuite échantillonné en un signal mono de huit bits à une fréquence de 8KHz.

C'est RANDOM.ORG qui publie ensuite sur son site :

<https://www.random.org/>

les données numériques obtenues.

Depuis le site RANDOM.ORG, on exporte 10 000 nombres décimaux compris dans l'intervalle [0; 1].

Ces données « naturelles » sont stockées dans le fichier nommé "random.txt".

Ces nombres sont répartis en deux colonnes comportant 5 000 nombres décimaux chacune.

**Question :** les nombres fournis permettent-ils de simuler un lancer de pièce équilibrée ?

**1. TRAITEMENT DES DONNÉES :**

**a.** Saisir les instructions suivantes :

```
In [1]: from math import *
import pandas as pa
```

```
In [2]: T=pa.read_csv('C:/.../random.txt', sep='\t', header=None)
```

```
In [3]: T.head()
```

Instruction qui permet d'afficher l'en-tête du tableau **T** associé à "random.txt"

Chemin d'accès indiquant l'emplacement de "Random.txt"  
**Les pointillés sont à compléter avec le professeur !!**

▷ '\t' indique à Python que les éléments de "random.txt" sont séparés par un espace  
▷ header=None indique à pandas que le fichier ne comporte pas de nom pour les colonnes.

**b. i.** Saisir l'instruction suivante et recopier les premières valeurs obtenues en sortie :

```
In [4]: T[0]
```

```
Out[4]:
```

**ii.** Saisir ensuite :

```
In [5]: X1=list(T[0])
```

Que représente la variable X1 ?

- iii. Obtenir la liste, nommée  $X_2$ , des valeurs contenues dans la deuxième colonne du fichier "random.txt".
  - iv. Construire la liste nommée  $L$ , des valeurs contenues dans le fichier "random.txt".
  - c. Écrire une fonction nommée **freq** ayant pour paramètres :
    - $r$  un nombre réel compris entre 0 et 1
    - $X$  une liste de valeurs.Cette fonction retournera la fréquence des valeurs de la liste  $X$  inférieures à  $r$ .
  - d. Déterminer la fréquence des valeurs de la liste  $L$  qui sont inférieures à 0,5.  
*Remarque : on nommera  $f$  la variable contenant cette fréquence.*
2. MODÉLISATION :
- a. Concevoir un modèle qui permet de simuler le lancer d'une pièce à partir de cette liste de valeurs  $L$ .
  - b. Pouvons-nous accepter ou non l'hypothèse suivante « la liste de valeurs de  $L$ , permet de simuler plusieurs lancers d'une pièce équilibrée » ?