♦ TP4 – Simulation d'expériences aléatoires

I. — Générateurs de nombres pseudo-aléatoires

À l'aide d'un ordinateur, il n'est pas possible de produire des nombres de façon totalement aléatoire. Cependant, certains algorithmes fournissent des suites de nombres suffisamment indépendants les uns des autres et qui ne semblent pas suivre de règles précises, de telle sorte qu'ils semblent aléatoires. On les appelle des nombres pseudo-aléatoires.

Dans toute la suite, on considère les nombres pseudo-aléatoires comme des nombres simulant des nombres choisis au hasard.

Pour obtenir des nombres pseudo-aléatoires en Python, il faut utiliser les fonctions du module `random`. Pour cela, on les importe à l'aide de l'instruction `from random import *`

Dans la suite, on utilisera les deux fonctions suivantes :

- `random()` renvoie un nombre flottant au hasard appartenant à l'intervalle $[0\,;1[$.
- si a et b sont deux entiers tels que $a \leqslant b$, alors `randint(a,b)` renvoie un nombre entier au hasard appartenant à $[\![a, b]\!]$. (On prendra garde au fait que, contrairement à l'habitude en Python, l'entier b est compris.)

II. — Simulation de quelques expériences aléatoires classiques

1) Cas équiprobable

Question 1. Écrire une fonction `lancer_de` qui simule un lancer de dé cubique équilibré, c'est-à-dire qui renvoie un entier choisi au hasard dans $[\![1, 6]\!]$.

Solution.

```
def lancer_de():
  return randint(1,6)
```

Question 2. Écrire une fonction `lancer_piece` qui simule un lancer de pièce équilibrée, c'est-à-dire qui renvoie `pile` ou `face` avec une probabilité $\frac{1}{2}$.

Solution.

```
def lancer_piece():
  if randint(0,1) == 0:
    return 'pile'
  else:
    return 'face'
```

2) Cas non équiprobable

On souhaite simuler le lancer d'une pièce truquée qui tombe sur *pile* avec une probabilité $p \in]0\,;1[$. Pour cela, on utilise le fait que l'instruction `random() < p` renvoie `True` avec une probabilité p.

Question 3. Écrire une fonction `lancer_piece_truquee` qui prend en argument un flottant `p` strictement compris entre 0 et 1 et qui renvoie `pile` avec une probabilité p et `face` avec une probabilité $1 - p$.

Solution.

```
def lancer_piece_truquee(p):
  if random() < p:
    return 'pile'
  else:
    return 'face'
```

III. — Estimation d'une probabilité à l'aide d'une fréquence

Question 4. Écrire une fonction `frequence_de_six` qui prend en argument un entier naturel non nul `n` et qui renvoie la fréquence d'apparition de 6 sur n simulations d'un lancer de dé équilibré. On rappelle que la fréquence d'apparition de 6 est le quotient du nombre de 6 obtenus par le nombre total de lancers.

Solution.

```
def frequence_de_six(n):
  c = 0
  for i in range(n):
    if lancer_de() == 6:
      c += 1
  return c/n
```

La loi faible des grands nombres, que nous verrons en fin d'année, assure que lorsqu'on considère un évènement E de probabilité p lors d'une expérience aléatoire alors la fréquence de réalisation de E lorsqu'on effectue un grand nombre de répétitions de l'expérience aléatoire est, en général, proche de p. Ainsi, si on lance un grand nombre de fois un pièce de monnaie équilibrée, alors la fréquence d'apparition de *pile* est, en général, proche de $\frac{1}{2}$.

On peut donc utiliser cela pour estimer la probabilité d'un évènement E en simulant un grand nombre de fois l'expérience aléatoire et en calculant la fréquence de réalisation de E.

Question 5. On lance deux dés équilibrés et on calcule la somme des nombres obtenus. On note E l'évènement « la somme est égale à 10 ».

1. Écrire une fonction `frequence_de_dix` qui prend en argument un entier naturel non nul `n` et qui renvoie la fréquence de réalisation de l'évènement E lorsqu'on simule n lancers de deux dés.

   Solution.

   ```
   def frequence_de_dix(n):
     c = 0
     for i in range(n):
       if lancer_de()+lancer_de() == 10:
         c += 1
     return c/n
   ```

2. En faisant plusieurs appels de la fonction `frequence_de_dix` avec $n = 100\,000$, estimer, à 10^{-3} près, la probabilité d'obtenir 10.

   Solution. En faisant 10 appels, on a trouvé 0.083665, 0.083734, 0.083734, 0.083426, 0.083289, 0.083261, 0.083344, 0.082912, 0.083332 et 0.083283. On peut donc estimer la probabilité de E à environ 0,083.

3. Démontrer que cette probabilité vaut $\frac{1}{12}$ et comparer avec l'estimation obtenue à la question précédente.

Solution. L'ensemble des lancers possibles est $[\![1,6]\!]^2$ qui est de cardinal $6^2 = 36$. Parmi ces lancers, ceux qui réalisent l'évènement E sont $(4,6)$, $(5,5)$ et $(6,4)$ donc, par équiprobabilité des lancers, $\mathtt{P}(E) = \frac{3}{36} = \frac{1}{12}$.

Sachant que $\frac{1}{12} \approx 0{,}8333$, ceci est cohérent avec l'estimation précédente.

IV. — Exercices

Exercice 1. Écrire une fonction `choix_liste` qui prend en argument une liste et qui renvoie un élément choisi au hasard dans cette liste.

Tester la fonction avec la liste `L = ['A', 'B', 'C', 'D', 'E']`.

Solution.

```
def choix_liste(L):
  k = randint(0, len(L)-1)
  return L[k]
```

Exercice 2. Écrire une fonction `choix_liste_avec_remise` qui prend en argument une liste `L` et un entier naturel non nul `n` et qui renvoie une liste constituée en choisissant successivement avec remise `n` éléments au hasard dans `L` (c'est-à-dire un même élément de la liste peut être choisi plusieurs fois).

Tester la fonction avec la liste `L = ['A', 'B', 'C', 'D', 'E']` et `n = 3`

Solution.

```
def choix_liste_avec_remise(L,n):
  M = []
  for i in range(n):
    M.append(choix_liste(L))
  return M
```

Exercice 3. Écrire une fonction `choix_liste_sans_remise` qui prend en argument une liste `L` et un entier naturel non nul `n` inférieur à `len(L)` et qui renvoie une liste constituée en choisissant successivement sans remise `n` éléments de `L` (c'est-à-dire un même élément de la liste ne peut pas être choisi plusieurs fois).

Indication. L'instruction `e = L.pop(k)` permet d'affecter `L[k]` à la variable `e` tout en supprimant cet élément de la liste `L`.

Tester la fonction avec la liste `L = ['A', 'B', 'C', 'D', 'E']` et `n = 3`

Solution.

```
def choix_liste_sans_remise(L,n):
  M = []
  for i in range(n):
    l = len(L)
    k = randint(0, l-1)
    e = L.pop(k)
    M.append(e)
  return M
```

Exercice 4. Une urne contient r boules rouges et b boules blanches. On tire simultanément n boules de l'urne avec $n \leqslant r+b$.

1. Écrire une fonction `tirages_simultanes` qui prend en arguments les trois entiers `r`, `b` et `n` et qui renvoie le nombre de boules rouges tirées.

   On pourra supposer que les boules rouges sont numérotées de 1 à r et les boules blanches de $r+1$ à $r+b$ et utiliser le fait que tirer n boules simultanément revient à tirer n boules successivement et sans remise.
2. Écrire une fonction `frequence_tirages_simultanes` qui prend en arguments les trois entiers `r`, `b` et `n` ainsi que deux autres entiers naturels `p` et `m` et qui renvoie la fréquence de réalisation de l'évènement S_p « on obtient p boules rouges » lorsqu'on effectue m tirages successives (en remettant dans l'urne les n boules tirées après chaque tirage).
3. On suppose ici que $r=15$, $b=8$, $n=5$ et $p=3$.
   a. Estimer la probabilité de S_3 à l'aide de la fonction `frequence_tirages_simultanes`.
   b. Calculer la valeur exacte de la probabilité de S_3 et comparer avec l'estimation trouvée précédemment.

Solution.

1.

```
def tirages_simultanes(r,b,n):
  L = ['r' for i in range(r)] + ['b' for i in range(b)]
  M = choix_liste_sans_remise(L, n)
  c = 0
  for e in M:
    if e == 'r':
      c += 1
  return c
```

2.

```
def frequence_tirages_simultanes(r,b,n,p,m):
  c = 0
  for i in range(m):
    if tirages_simultanes(r,b,n) == p:
      c += 1
  return c/m
```

3. a. En appelant la fonction `frequence_tirages_simultanes` avec $r=15$, $b=8$, $n=5$, $p=3$ et $m=100000$, on trouve une fréquence proche de 0,38.

b. Le nombre total de tirages possibles est $\binom{23}{5} = 33649$ et le nombre de tirages contenant 3 boules rouges est $\binom{15}{3} \times \binom{8}{2} = 12740$ donc, par équiprobabilité, $\mathbf{P}(S_3) = \frac{12740}{33649} = \frac{1820}{4807} \approx 0{,}379$ ce qui est en accord avec l'estimation précédente.

Exercice 5. On considère des chaînes de n caractères choisis parmi `'C'`, `'G'`, `'T'` et `'A'`. Pour $n=6$, on a, par exemple, les chaînes suivantes `'CTTACT'`, `'CGATCAA'` et `'GTGCGA'`

1. Écrire une fonction `chaine` qui prend en argument l'entier `n` et qui renvoie une chaîne aléatoire respectant le format décrit ci-dessus.
2. Écrire une fonction `frequence_AT` qui prend en arguments deux entiers et `n` et `m` et qui calcule la fréquence de chaînes commençant par `'AT'` parmi `m` chaînes de longueurs `'n'` choisies aléatoirement.

3. Utiliser la fonction `frequence_AT` pour estimer la probabilité p qu'une chaîne de longueur 6 choisie aléatoirement commence par `'AT'`.
4. Calculer la valeur exacte de p et comparer avec l'estimation obtenue à la question précédente.

Solution.

1.

```
def chaine(n):
  L = ['C', 'G', 'T', 'A']
  ch = ''
  for i in range(n):
    k = randint(0, 3)
    ch += L[k]
  return ch
```

2.

```
def frequence_AT(n,m):
  c = 0
  for i in range(m):
    ch = chaine(n)
    if ch[:2] == 'AT':
      c += 1
  return c/m
```

3. En appellant la fonction `frequence_AT` avec $n = 6$ et $m = 100000$, on trouve une fréquence proche de 0,06.

4. Il y a 4^6 chaînes possibles et, parmi celles-ci, 4^4 commence par `'AT'`. Par équiprobabilité, on en déduit que $p = \frac{4^4}{4^6} = \frac{1}{4^2}$ i.e. $p = \frac{1}{16} \approx 0{,}0625$, ce qui est cohérent avec l'approximation précédente.

Exercice 6. On effectue une succession infinie de lancers d'une pièce truquée qui amène `pile` avec une probabilité $p \in]0\,;1[$.

1. Écrire une fonction `deuxieme_pile` qui prend en argument un flottant `p` compris entre 0 et 1 et qui simule ces lancers successifs et qui renvoie le rang d'apparation du deuxième `pile`.
2. Écrire une fonction `frequence_deuxieme_pile` qui prend en arguments un flottant `p` compris entre 0 et 1 et deux entiers naturels non nuls `n` et `r` et qui renvoie la fréquence de l'évènement S_r : « le deuxième *pile* apparaît lors du r-ème lancer » lorsqu'on simule n successions de lancers de pièce.
3. Utiliser la fonction `frequence_deuxieme_pile` pour estimer la probabilité de S_r et comparer avec la valeur exacte qui est $\mathbf{P}(S_r) = (r-1)p^2(1-p)^{r-2}$.

Solution.

1.

```
def deuxieme_pile(p):
    c = 0
    k = 0
    while c<2:
        k += 1
        if lancer_piece_truquee(p) == 'pile':
            c += 1
    return k
```

2.

```
def frequence_deuxieme_pile(p,n,r):
    c = 0
    for i in range(n):
        k = deuxieme_pile(p)
        if k == r:
            c += 1
    return c/n
```

3. En prenant $p = 0{,}3$, $n = 100000$ et $r = 5$, on trouve une fréquence proche de 0,12 et $4 \times 0{,}3^2 \times 0{,}7^3 = 0{,}12348$, en prenant $p = 0{,}6$, $n = 100000$ et $r = 7$, on trouve une fréquence proche de 0,02 et $6 \times 0{,}6^2 \times 0{,}4^5 = 0{,}0221184$ et en prenant $p = 0{,}1$, $n = 100000$ et $r = 10$, on trouve une fréquence proche de 0,04 et $9 \times 0{,}1^2 \times 0{,}9^8 = 0.0387420489$. Ainsi, les estimations sont proches des probabilités.

Exercice 7. Une urne contient une boule blanche et une boule rouge.

On tire dans cette urne une boule, on note sa couleur et on la remet dans l'urne accompagnée de deux autres boules de la même couleur puis on répète l'opération.

On a vu dans l'exercice 4 du chapitre 4 que la probabilité de n'obtenir que des boules rouges lors des n premiers tirages est $p_n = \frac{1}{2} \times \frac{3}{4} \times \cdots \times \frac{2n-1}{2n}$.

1. Écrire une fonction `tirage` qui prend en argument un entier naturel non nul `n` et qui renvoie le nombre de boules rouges tirées lorsque de la simulation de n tirages.
2. Écrire une fonction `frequence_rouge` qui prend en arguments deux entiers naturels non nuls `n` et `m` et qui renvoie la fréquence de l'évènement « on n'a tiré que des boules rouges lors des n premiers tirages » lorsqu'on simule m fois l'expérience.
3. Écrire une fonction `probabilite` qui prend en argument un entier naturel non nul `n` et qui renvoie le résultat du calcul de p_n.
4. Comparer les résultats renvoyés par les fonctions `frequence_rouge` et `probabilite` pour différentes valeurs de `n` et pour de grandes valeurs de `m`.

Solution.

1.

```
def tirage(n):
  c = 0
  L = ['r', 'b']
  for i in range(n):
    k = randint(0, len(L)-1)
    if L[k] == 'r':
      c +=1
      L += ['r', 'r']
    else:
      L += ['b', 'b']
  return c
```

2.

```
def frequence_rouge(n,m):
  c = 0
  for k in range(m):
    nb = tirage(n)
    if nb == n:
      c += 1
  return c/m
```

3.

```
def probabilite(n):
  p = 1
  for k in range(1,n+1):
    p = p * (2*k-1)/(2*k)
  return p
```

4. Voici les résultats de plusieurs essais :

```
print(frequence_rouge(5,100000), probabilite(5))
> 0.24763 0.24609375

print(frequence_rouge(12,100000), probabilite(12))
> 0.16039 0.1611802577972412

print(frequence_rouge(7,1000000), probabilite(7))
> 0.20941 0.20947265625

print(frequence_rouge(15,1000000), probabilite(15))
> 0.144579 0.14446444809436798
```

Exercice 8. On lance un dé cubique équilibré jusqu'à obtenir 6 pour la première fois. On s'intéresse à l'évènement A : « On n'a obtenu que des nombres pairs ».

1. Écrire les fonctions nécessaires pour estimer la probabilité de A.
2. Comparer avec la valeur exacte de cette probabilité (voir Chapitre 4, Exercice 14).

Solution.

1. On écrit une première fonction `tous_pairs` qui renvoie `True` si on n'a obtenu que des nombres pairs jusqu'au premier 6 et `False` sinon.

```
def tous_pairs():
  c = 0
  p = 0
  f = 0
  while f != 6:
    c += 1
    f = lancer_de()
    if f%2 == 0:
      p += 1
  return p == c
```

   On écrit une seconde fonction `frequence_tous_pairs` qui renvoie la fréquence de réalisation de l'évènement « N'obtenir que des nombres pairs jusqu'au premier 6 » sur n simulations.

```
def frequence_tous_pairs(n):
  c = 0
  for i in range(n):
    if tous_pairs:
      c += 1
  return c/n
```

2. En appelant la fonction `frequence_tous_pairs` avec $n = 100000$, on obtient des valeurs très proches de 0,25 ce qui est en accord avec la probabilité de $\frac{1}{4}$ calculée dans l'exercice 14 du chapitre 4.