Simulations Probabilistes

1 Concept

1.1 Définitions

▲ DÉFINITION 1

Simuler un tirage dans l'espace probabilisé $(\Omega, \mathbb{P})$, c'est écrire une fonction Python qui a les caractéristiques suivantes :

1. Lorsqu'on l'éxécute, la fonction retourne un élément ω de Ω
2. Lorsqu'on l'éxécute un grand nombre de fois, la fonction renvoie chaque objet ω de Ω dans un pourcentage de cas "égal" à $\mathbb{P}(\{\omega\})$.
3. Le nombre et les résultats de ses exécutions précédentes ne sont pas pris en compte lors d'un appel de la fonction.

▲ DÉFINITION 2

Soit X une variable aléatoire de loi L. **Simuler** la variable aléatoire X, ou la loi L, c'est écrire une fonction Python telle que :

— cette fonction tire au hasard un réel, et renvoie un certain objet en sortie en fonction de ce résultat

— La loi de l'objet renvoyé, est exactement la loi L.

— Cette fonction est indépendante de X.

L'objet renvoyé à chaque fois en sortie est aussi appelé une **réalisation** de la loi L ou de la variable aléatoire X.

1.2 Exemples en Python – fonctions de base

Nous utiliserons le module numpy.random

```
import numpy as np

import numpy.random as rd
```

1.2.1 Tirage uniforme sur un intervalle d'entiers

L'expression randint(a, b) permet de choisir un entier au hasard dans l'intervalle $[\![a, b[\![$.

— Cela correspond à l'univers $([\![a, b-1]\!], \mathbb{P})$ où $\mathbb{P}$ est la probabilité uniforme.

— Si X désigne la variable aléatoire correspondant au résultat de la fonction randint(a, b), alors pour tout entier k dans $[\![a, b-1]\!]$, on a $\mathbb{P}(X = k) = \dfrac{1}{b-a}$

Exemple

```
[rd.randint(10) for _ in range(20)]
> [5, 1, 8, 7, 8, 8, 9, 9, 5, 1, 7, 6, 3, 6, 6, 4, 0, 1, 7, 8]

[rd.randint(8,10) for _ in range(20)]
> [8, 8, 9, 8, 9, 8, 9, 8, 8, 8, 8, 9, 8, 9, 8, 8, 8, 9, 8, 9]
```

Exercice 1 Donner un morceau de code qui renvoie une matrice de taille (3,4) contenant des résultats de simulations lancers de dé.

```
[[rd.randint(1,7) for _ in range(4)] for _ in range(3)]
>[[1, 2, 3, 4], [2, 5, 6, 3], [4, 6, 6, 5]]
```

1.2.2 Tirage uniforme sur $[0, 1]$

La fonction random tire un réel compris dans l'intervalle $[0, 1[$.

— Cela correspond à l'univers $([0, 1[, \mathbb{P})$ où $\mathbb{P}$ est la probabilité uniforme.

— Si X désigne la variable aléatoire correspondant au résultat de la fonction random, alors pour tout a et b dans $[0, 1]$ avec $a < b$, on a $\mathbb{P}(a \leqslant X \leqslant b) = b - a$

Exemple

```
rd.random()
>0.9734582368079974


[rd.random() for _ in range(20)]
>[0.6995211719968324, 0.8862245956395243, 0.9873494779676141, 0.6126809499045885,
    0.8853255442652447, 0.5401606412945948, 0.9360697298863336, 0.17426606452207283,
    0.20815648982363855, 0.9789902327962481, 0.3481946730853992, 0.7828191093575398,
    0.991094019694116, 0.23840423144429856, 0.9106685047231456, 0.09693740362321013,
    0.46800438143054535, 0.45863908128518927, 0.606330960066964, 0.9578026278799725]

sum([rd.random() for _ in range(20)])/20
>0.49478671950937636
```

Exercice 2 Que fait la fonction Python suivante ?

```
    def neo_random():
    x=0
    for k in range(1,31):
        x+=rd.randint(0,10)/(10**k)
    return x
```

Grâce à la fonction random, programmer une fonction dont l'appel a le même effet que :

1. la fonction randint(0,2) ;
2. la fonction randint(0, 10).

Réponses :

1. Elle a le même effet que randint
2.

```
def randint01():
    return 1*(rd.random()>0.5)

def randint09():
    return floor(10*rd.random())
```

Exercice 3 $\boxed{\text{Classique}}$ Dans la classe de PCX, quelle est la probabilité que deux d'entre vous soient nés le même jour ?

Pour information la solution mathématique

Cette solution est écrite en négligeant les années bissextiles ; et pour r personnes ; on va évaluer la probabilité de A=« les r personnes ont leur anniversaire à des dates deux à deux distinctes ».

L'univers Ω est l'ensemble $[\![1,365]\!]^r$, l'événement A est l'ensemble des arrangements à r éléments appartenant à Ω.

On fait l'hypothèse que la date de naissance de chaque personne est uniformément répartie sur les 365 jours de l'année (probabilité uniforme) et que les « choix »des dates anniversaires des différentes personnes sont indépendants, donc on a probabilité uniforme sur Ω. Le nombre total de choix pour les r personnes est alors 365^r. Le cardinale de A est A_{365}^r et la probabilité cherchée est $\dfrac{A_{365}^r}{365^r}$. Cette probabilité est nulle si $r > 365$ et sinon, est égale à $\dfrac{365!}{(365-r)!365^r}$.

Application numérique pour $r = 40$, cette probabilité est égale à $0{,}11$.

Ainsi la probabilité de l'événement contraire (recherchée) est 89% **.**

1.3 Simulations de lois classiques du cours

Loi binomiale (exemple typique de simulation algébrique)
Loi géométrique
Un exercice mêlant les deux simulations
Un exercice mêlant du discret et du continu
Comment faire pour tirer uniformément dans un ensemble qui n'est pas un ensemble d'entiers ?
Comment tirer selon une loi donnée, dans un ensemble de nombre réels ?

Exercice 4 On se donne $2n$ urnes numérotées de 1 à $2n$. L'urne numérotée i contient i jetons, numérotés de 1 à i. On choisit une urne au hasard. Dans l'urne choisie, on tire un jeton au hasard.
On note p_n la probabilité que ce soit le jeton n qui est tiré. Conjecturer informatiquement un équivalent de p_n (lorsque $n \to +\infty$).

2 Simulations de variables aléatoires non indépendantes

2.1 Réduction des possibilités

Exemple

Ecrire une fonction qui reçoit un entier naturel n et un entier $k \leqslant n$ en entrée, et qui renvoie en sortie une liste de longueur k, formé de k nombres entiers **distincts** tirés au hasard entre 1 et n (chaque coordonnée suivra une loi uniforme).

```python
def liste_entiers_distincts(n,k):
    entiers_restants=[i for i in range(1,n+1)]
    L=[]
    while k>0:
        indice=rd.randint(len(entiers_restants))
        i=entiers_restants[indice]
        del(entiers_restants[indice])
        k-=1
        L.append(i)
    return L
```

Exemple

Donner une fonction qui prend en arguments trois nombres entiers n, p et N avec $N \leqslant n \times p$, et qui renvoie en sortie une matrice M de taille $n \times p$, dont N coefficients valent 1 et $np - N$ coefficients valent 0. Chaque matrice aura la même probabilité.

2.2 Connaissance de la loi conjointe d'un couple

Même principe en simulant sur une matrice

3 Espérance et Loi des grands nombres

Pour les simulations on va plutôt utiliser :

T THÉORÈME 1 — INÉGALITÉ DE BIENAYMÉ-TCHEBYCHEV

Toute variable aléatoire réelle X vérifie l'inégalité :

$$\forall \varepsilon > 0 \ \mathbb{P}\big(|X - \mathbb{E}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Et surtout le résultat :

Exercice 5 $\boxed{\text{Classique}}$ Soit $(X_1, \dots, X_n)$ des variables aléatoires réelles indépendantes et de même loi. On note $m = E(X_1)$.
Montrer que, pour tout $\epsilon > 0$, on a

$$\lim_n P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m\right| \geqslant \epsilon\right) = 0$$

`#Loi faible des grands nombres #Loi des grands nombres #LGN` **Solution :** On pose $Y = \frac{1}{n}\sum_{i=1}^{n} X_i$, on lui applique l'inégalité de B-Tchebitchev; et on se rend compte que la probabilité de l'énoncé est donc majorée par $\dfrac{V(Y)}{\epsilon^2} = \dfrac{V(X_1)}{n\epsilon^2}$ puisque les X_i ont tous la même loi et sont indépendantes. Le théorème des gendarmes conclut.

Exercice 6 Retrouver par simulation l'espérance de lois classiques.

Exercice 7 On dispose de 100 poulets assis en cercle; chaque poulet donne un coup de bec a son voisin de droite ou à son voisin de gauche, au hasard. On appelle N le nombre de poulet qui n'ont pas reçu de coup de bec. Calculer $E(N)$.
Solution mathématique Appelons X_i la variable aléatoire qui vaut 1 si le poulet i recoit un c de bec,n et 0 sinon. On

a $X_i \sim B(1/4)$ et $N = X_1 + \cdots + X_{100}$. Ainsi par linéarité, on a $E(N) = \cdots = 25$.

Exercice 8 On se donne $2n$ urnes numérotées de 1 à $2n$. L'urne numérotée i contient i jetons, numérotés de 1 à i. On choisit une urne au hasard. Dans l'urne choisie, on tire un jeton au hasard.
Donner l'espérance du numéro de jeton (par simulation informatique).

4 Exercices récapitulatifs du cours

1. Ecrire une fonction qui reçoit un entier naturel n en argument, et qui tire au hasard un nombre entier compris entre 1 et n (selon une loi uniforme).
2. Ecrire une fonction qui reçoit un entier naturel n et un entier $k \leqslant n$ en entrée, et qui renvoie en sortie une liste de longueur k, formés de k nombres entiers tirés au hasard entre 1 et n (selon une loi uniforme).
3. Donner dans chaque cas une fonction qui renvoie sous forme de liste :
   (a) N réalisations indépendantes d'une loi de Bernoulli de paramètre p,
   (b) N réalisations indépendantes d'une loi binomiale de paramètres n et p,
   (c) N réalisations indépendantes d'une loi géométrique de paramètre p (programme de 2nde année)
   (d) N réalisations indépendantes d'une loi uniforme sur $\{a_1, \ldots, a_n\}$
   (e) N réalisations indépendantes d'une variable aléatoire X suivant une loi discrète sur $\{a_1, \ldots, a_n\}$, de loi de probabilité donnée par $P(X = a_k) = p_k$ pour tout $k \in \{1, \ldots, n\}$
   (f) N réalisations indépendantes d'une loi de Poisson de paramètre λ (programme de 2nde année)
4. Ecrire une fonction qui reçoit un entier naturel n et un entier $k \leqslant n$ en entrée, et qui renvoie en sortie une liste de longueur k, formé de k nombres entiers **distincts** tirés au hasard entre 1 et n (chaque coordonnée suivra une loi uniforme).

Exercice 9 Soit des variables aléatoires $X_{i,j} \sim \mathcal{B}(p)$, indépendantes, et $X :=$ la matrice carrée de taille $n \times n$ dont les coefficients sont les $X_{i,j}$.

1. Si A est une matrice à coefficients dans $\{0,1\}$, donner la probabilité d'avoir $X = A$ en fonction des coefficients de A, et de p.
2. Donner la loi de $\operatorname{Tr} X$
3. Donner la probabilité que chaque ligne de X ait la somme s
4. Donner la probabilité que chaque ligne de X ait la même somme.
5. Donner la probabilité que chaque ligne de X et chaque colonne de X ait la somme $n-1$.

Exercice 10 $\boxed{\text{TD}}$ On considère une urne contenant n jetons numérotés de 1 à n ($n \geqslant 2$). On prélève ces jetons au hasard, un par un et sans remise. On note $(u_1, u_2, \ldots, u_n)$ la liste des numéros tirés.
Pour $2 \leqslant i \leqslant n$, on dit qu'il y a record à l'instant i si $u_i > \max(u_1, \ldots, u_{i-1})$. On convient qu'il y a systématiquement record à l'instant 1.

1. Calculer, pour $1 \leqslant i \leqslant n$, la probabilité r_i qu'il y ait record à l'instant i.
2. Calculer les probabilités que, durant la totalité des tirages, on assiste exactement à :
   (a) un seul record ; (b) n records ; (c) deux records.

Solution math On représente un résultat de l'expérience par un n-uplet $(u_1, u_2, \ldots, u_n)$ qui s'identifie à la permutation de $[\![1, n]\!]$ qui à i associe u_i. On a donc $\operatorname{card}(\Omega) = n!$. On fait l'hypothèse que tous les tirages sont équiprobables.

1. Il y a record à l'instant i si $u_i > \max(u_1, \ldots, u_{i-1})$. L'ensemble $\{u_1, \ldots, u_i\}$ étant choisi, il faut que le plus grand élément soit placé au i-ème tirage, l'ordre des $i-1$ autre éléments étant indifférent ; par ailleurs, l'ordre dans lequel sont tirés les $n-i$ jetons restant est également indifférent. On obtient :
$$r_i = \frac{\binom{n}{i}(i-1)!(n-i)!}{n!} = \frac{1}{i}.$$
Au voit que le résultat est valable pour $i = 1$ car, au premier tirage, il y toujours record.
2. (a) — S'il y a un seul record, il est obtenu au premier tirage. On peut remarquer que lorsqu'on tire la boule numéro n, il y a record. Elle doit donc être tirée en premier.
   — Réciproquement, si l'on tire la boule n en premier, il n'y a qu'un record car, pour $i \geqslant 2$, on a $\max(u_1, \ldots, u_{i-1}) = n \geqslant u_i$.
   Il faut donc tirer la boule n en premier et les autres dans un ordre quelconque. La probabilité d'avoir un seul record est $p_1 = \dfrac{(n-1)!}{n!} = \dfrac{1}{n}$.
   (b) S'il y a deux records, l'un est obtenu au rang 1. Déterminons la probabilité q_i d'obtenir deux records, aux rangs 1 et i, où $i \geqslant 2$.
   — Une condition nécessaire est $u_i = n$.

— Réciproquement, si cette condition est réalisée, il ne peut pas y avoir de record aux rangs $i+1,\ldots, n$. Pour qu'il n'y ait pas de record entre les rangs 2 et $i-1$, il faut et il suffit que $u_1 = \max(u_2,\ldots,u_{i-1})$.

On raisonne comme dans la question 1. On choisit $\{u_1,\ldots,u_{i-1}\}$, sous-ensemble de $[\![1, n-1]\!]$; le plus grand élément de cet ensemble est placé en premier, les autres dans un ordre quelconque. Pour u_i, il n'y a pas de choix et les $n-i$ éléments restant sont tirés dans un ordre quelconque. On obtient :

$$q_i = \frac{\binom{n-1}{i-1}(i-2)!(n-i)!}{n!} = \frac{1}{n(i-1)}.$$

La probabilité qu'il y ait exactement deux records est :

$$p_2 = \sum_{i=2}^{n} q_i = \sum_{i=2}^{n} \frac{1}{n(i-1)} = \frac{1}{n}\sum_{k=1}^{n-1}\frac{1}{k}.$$

(c) S'il y a n records, on a pour $1 \leqslant i \leqslant n$, $u_i > \max(u_1,\ldots,u_{i-1}) \geqslant u_{i-1}$. La suite $(u_i)_{1\leqslant i\leqslant n}$ est strictement croissante et $u_i = i$ pour tout i. Il est clair, réciproquement que si cette condition est réalisée, il y a n records. Une seule permutation convient, donc la probabilité d'obtenir n records est $p_n = \dfrac{1}{n!}$.

Exercice 11 Soit $p \in]0,1[$.
On considère une population de cellules. A chaque génération chacune des cellules, indépendamment des autres, ou bien se divise en deux avec probabilité p, ou bien meurt avec probabilité $1-p$. On note X_n le nombre de cellules à la n eme génération. La génération 0 est formée de $X_0 = 1$ cellule. On appelle enfin d_n la probabilité qu'il n'y ait aucune cellule vivante à la génération n.
Quelle est la probabilité que la population s'éteigne un jour ? on discutera selon les valeurs de p.