Correction - DS 10

Devoir surveillé du Mardi 4 Mars

Exercice 1

1. On peut utiliser le code suivant :

```
def exponentielle(lambda):
      u = rd.random()
      return(-(1/lambda)*np.log(1-u))
```

2. (a) On résout : $p = 1 - e^{-\lambda} \Leftrightarrow \lambda = -\ln(1-p)$. Avec le résultat admis, et en se souvenant que `np.floor` est la fonction partie entière sur `Python`, on propose la fonction suivante :

```
def geom(p):
      lambda = - np.log(1-p)
      x = exponentielle(lambda)
      return(np.floor(x)+1)
```

(b) On procède ainsi :

```
x = np.zeros(10000) #vecteur de taille 10000 avec que des 0
for k in range(10000):
     x[k] = geom(0.2) #remplace la k-ème composante par une
simulation
```

(c) La commande `np.mean(x)` calcule la moyenne des composantes du vecteur `x`. Par la loi faible des grands nombres, elle devrait être proche de l'espérance $\frac{1}{p} = \frac{1}{0.2} = 5$. C'est bien ce qu'on obtient ici.

De même, `np.mean((x-m)**2)` calcule la variance des composantes du vecteur `x`. Par (un corollaire de) la loi faible des grands nombres, elle devrait être proche de la variance théorique $\frac{q}{p^2} = \frac{0.8}{0.04} = 20$. C'est bien là aussi ce qu'on observe.

(d) Le premier graphique permet de réprésenter l'histogramme associé à la série définie par le vecteur `x` (créé à la question 2.(b)), c'est-à-dire le vecteur de 10000 réalisation de la variable aléatoire Y, pour les classes définie par le vecteur `c`, c'est-à-dire les 11 classes $[0.5, 1.5], [1.5, 2.5], \ldots, [10.5, 11.5]$, respectivement centrées en $1, 2, \ldots, 11$. Cet histogramme permet de visualiser la loi de Y.

Le deuxième graphique permet de visualiser la loi d'une variable aléatoire géométrique de paramètre 0.2 (en faisant 10000 simulations de cette loi et en affichant l'histogramme des fréquences qui sont alors proches des probabilités théoriques avec la loi faible des grands nombres).

On constate que ces deux graphiques sont quasi identiques, signe que les lois sont très proches, voire les mêmes. Ceci confirme bien que la variable Y ainsi définie dans l'énoncé, suit bien une loi géométrique.

Exercice 2

1. Sachant $(X = k)$, Y compte le nombre de succès au cours de k épreuves de Bernoulli indépendantes de même paramètre $p = 0,05$. Donc Y suit une loi binomiale de paramètres k et $0,05$.

2. On utilise la question précédente pour compléter la fonction :

```
def simulY():
    x = rd.poisson(20)
    y = rd.binomial(x, 0.05)
    return(y)
```

3. Avec une boucle `for`.

```
def SimulY(N):
    M = np.zeros(N)
    for k in range(N):
        M[k] = simulY()
    return(M)
```

4. Voici la fonction `loipoisson` :

```
def loipoisson(lb)
    V = np.zeros(11)
    V[0] = np.exp(-lb)
    for k in range(1,11):
        V[k] = V[k-1]*lb/k
    return(V)
```

5. La variable `U` est un vecteur ligne contenant un échantillon de 100000 simulations de la variable aléatoire Y.

   La variable `c` est un vecteur ligne contenant les classes pour le tracé de l'histogramme (de -0.5 à 0.5 pour la modalité 0, de 0.5 à 1.5 pour la modalité 1, ..., de 8.5 à 9.5 pour la modalité 9).

   La variable `n` est un vecteur ligne contenant les entiers de 0 à 9.

   Enfin, la variable `V` est un vecteur ligne contenant les 10 premières probabilités théoriques d'une loi de Poisson de paramètre 1.

   Ce programme permet de tracer deux diagrammes en bâtons :

   - Le graphe de gauche représente le diagramme en bâtons des fréquences de l'échantillon contenant les 100000 simulations de Y. Le bâton d'abscisse i indique en ordonnée la fréquence d'apparition de i dans l'échantillon généré.
   - Le graphe de droite représente le diagramme en bâtons des 10 premières probabilités théoriques d'une loi de Poisson de paramètre 1. Le bâton d'abscisse $i \in [\![0, 9]\!]$ indique en ordonnée la probabilité $\dfrac{e^{-1}}{i!}$.

6. Par comparaison des deux diagrammes obtenus, on constate que les fréquences empiriques de notre échantillon correspondant approximativement aux probabilités théoriques. D'après la loi forte des grands nombres, on peut donc supposer que Y suit une loi de Poisson de paramètre 1.

   C'est effectivement le cas : nous avons démontré dans l'exercice 15 du TD 9 que Y suit une loi de Poisson de paramètre $\lambda p = 20 \times 0.05 = 1$.

Exercice 3

1. (a) Sur $X(\Omega) =]\lambda; +\infty[$, $F(x) = 1 - \dfrac{\lambda^k}{x^k}$ donc F est dérivable et $F'(x) = \dfrac{k\lambda^k}{x^{k+1}} > 0$.
F est ainsi continue et strictement croissante sur $]\lambda; +\infty[$. Elle réalise donc une bijection de $]\lambda; +\infty[$ dans $]0; 1[$.

Soit $y \in]0; 1[$ fixé. Résolvons l'équation $F(x) = y$ d'inconnue $x \in]\lambda; +\infty[$:

$$F(x) = y \Leftrightarrow 1 - \frac{\lambda^k}{x^k} = y \Leftrightarrow \frac{\lambda^k}{x^k} = 1 - y \Leftrightarrow x^k = \frac{\lambda^k}{1-y} \Leftrightarrow x = \frac{\lambda}{\sqrt[k]{1-y}}.$$

Ainsi, $\forall y \in]0; 1[$, $F^{-1}(y) = \dfrac{\lambda}{\sqrt[k]{1-y}}$.

(b) On en déduit alors la fonction suivante pour simuler une loi de Pareto de paramètres λ et k avec la méthode d'inversion :

```
def Pareto1(lbd, k):
    U = rd.random()
    X = lbd/(1-U)**(1/k)
    return(X)
```

2. (a) Pour tout i, $X_i(\Omega) =]0, 1]$ donc $(\max(X_1, \ldots, X_k))(\Omega) =]0, 1]$ et donc $Y(\Omega) = [\lambda, +\infty[$.
Donc $F_Y(x) = 0$ si $x < \lambda$.
Si $x \geq \lambda$, on a :

$$\begin{aligned}
F_Y(x) &= P(Y \leq x) = P\left(\frac{\lambda}{x} \leq \max(X_1, \ldots, X_k)\right) \\
&= 1 - P\left(\max(X_1, \ldots, X_k) < \frac{\lambda}{x}\right) = 1 - P\left(\bigcap_{i=1}^{k} (X_i < \frac{\lambda}{x})\right) \\
&= 1 - \prod_{i=1}^{k} P(X_i < \frac{\lambda}{x}) \quad \text{(par indépendance des } X_i\text{)} \\
&= 1 - \prod_{i=1}^{k} P(X_i \leq \frac{\lambda}{x}) \quad \text{(car les } X_i \text{ sont à densité)} \\
&= 1 - \prod_{i=1}^{k} \frac{\lambda}{x} \quad \text{(car } X_i \hookrightarrow \mathcal{U}(]0, 1]) \text{ et } \frac{\lambda}{x} \in]0, 1]) \\
&= 1 - \frac{\lambda^k}{x^k}.
\end{aligned}$$

Donc Y suit une loi de Pareto de paramètres λ et k.

(b) On en déduit le programme suivant :

```
def Pareto2(lbd, k):
    U = rd.random(k)
    Y = lbd/np.max(U)
    return(Y)
```

Exercice 4

1. La variable `U` est une matrice ligne de 100 termes telle que, pour tout $1 \leq k \leq 100$, $U(k) = k$.

   La variable `V` est une matrice ligne de 100 termes telle que, pour tout $1 \leq k \leq 100$, $V(k) = \sum_{i=1}^{k} \frac{1}{i} = H_k$ (c'est la somme cumulée des inverses des entiers contenus dans la matrice ligne `U`).

   Ce programme trace les 100 premiers termes de la suite des sommes partielles $(H_n)_{n\geq 1}$ de la série harmonique.

   On constate que cette suite des sommes partielles divergent vers $+\infty$. Ceci est conforme aux résultats du cours puisque c'est une série de Riemann de paramètre $\alpha = 1 \leq 1$.

2. Voici la fonction `seuil` demandée :

```
def seuil() :
    N=1
    H=1
    while H<10 :
        N = N+1
        H = H+1/N
    return(N)
```

   Après avoir exécuté la fonction, on obtient $N = 12367$. La série diverge donc très lentement vers $+\infty$.

3. La variable `U` est une matrice ligne de 99 termes telle que, pour tout $2 \leq k \leq 100$, $U(k-1) = k$.

   La variable `V` est une matrice ligne de 99 termes telle que, pour tout $2 \leq k \leq 100$, $V(k-1) = 1 + \sum_{i=2}^{k} \frac{1}{i} = H_k$.

   La variable `W` est une matrice ligne de 99 termes telle que, pour tout $2 \leq k \leq 100$, $W(k-1) = \ln(k)$.

   La variable `T` est une matrice ligne de 99 termes (car on commence au deuxième terme pour ne pas diviser par 0) telle que, pour tout $2 \leq k \leq 100$, $T(k-1) = \dfrac{H_k}{\ln(k)}$.

   Ce programme trace les 99 premiers termes de la suite $\left(\dfrac{H_n}{\ln(n)}\right)_{n\geq 2}$.

   On remarque que la suite $\left(\dfrac{H_n}{\ln(n)}\right)_{n\geq 2}$ semble converger vers 1. En d'autres termes, on a donc $H_n \sim \ln(n)$. Ceci avait été démontré dans l'exercice 1 de la séance d'approfondissement 4.

4. (a) Soit $n \in \mathbb{N}^*$. Par décroissance de la fonction inverse sur $\mathbb{R}_+^*$ puis en intégrant sur des bornes croissantes, on a :

$$n \leq t \leq n+1 \Rightarrow \frac{1}{n} \geq \frac{1}{t} \geq \frac{1}{n+1} \Rightarrow \int_n^{n+1} \frac{1}{n} dt \geq \int_n^{n+1} \frac{1}{t} dt \geq \int_n^{n+1} \frac{1}{n+1} dt.$$

   Comme $\displaystyle\int_n^{n+1} \frac{1}{n} dt = \frac{1}{n} \int_n^{n+1} 1 dt = \frac{1}{n} [t]_n^{n+1} = \frac{1}{n}$ et de même $\displaystyle\int_n^{n+1} \frac{1}{n+1} dt = \frac{1}{n+1}$, on obtient finalement que :

$$\frac{1}{n+1} \leq \int_n^{n+1} \frac{1}{t} dt \leq \frac{1}{n}.$$

(b)
- $(u_n)_{n\geq 1}$ est décroissante :

$$\begin{aligned} u_{n+1}-u_n &= \left(\sum_{k=1}^{n+1}\frac{1}{k}-\ln(n+1)\right)-\left(\sum_{k=1}^{n}\frac{1}{k}-\ln(n)\right) \\ &= \frac{1}{n+1}-\ln(n+1)+\ln(n)=\frac{1}{n+1}-\int_n^{n+1}\frac{1}{t}dt\leq 0, \end{aligned}$$

d'après la question précédente.
- $(v_n)_{n\geq 1}$ est croissante :

$$\begin{aligned} v_{n+1}-v_n &= \left(\sum_{k=1}^{n+1}\frac{1}{k}-\ln(n+2)\right)-\left(\sum_{k=1}^{n}\frac{1}{k}-\ln(n+1)\right) \\ &= \frac{1}{n+1}-\ln(n+2)+\ln(n+1)=\frac{1}{n+1}-\int_{n+1}^{n+2}\frac{1}{t}dt\geq 0, \end{aligned}$$

d'après la question précédente.
- On a :

$$\begin{aligned} v_n-u_n &= \left(\sum_{k=1}^{n}\frac{1}{k}-\ln(n+1)\right)-\left(\sum_{k=1}^{n}\frac{1}{k}-\ln(n)\right)=-\ln(n+1)+\ln(n) \\ &= \ln\left(\frac{n}{n+1}\right)=\ln\left(\frac{1}{1+\frac{1}{n}}\right)\underset{n\to+\infty}{\longrightarrow}0 \end{aligned}$$

par continuité de ln en 1.

On a démontré que $(u_n)_{n\geq 1}$ est décroissante, que $(v_n)_{n\geq 1}$ est croissante et que $\lim\limits_{n\to+\infty} v_n - u_n = 0$. Donc $(u_n)_{n\geq 1}$ et $(v_n)_{n\geq 1}$ sont adjacentes.
D'après le théorème des suites adjacentes, $(u_n)_{n\geq 1}$ et $(v_n)_{n\geq 1}$ convergent vers une même limite γ.

(c) Voici la fonction `gamma` demandée :

```
def gamma(eps) :
    n = 1
    H = 1
    u = H
    v = H-np.log(2)
    while np.abs(u-v)>eps :
        n = n+1
        H = H+1/n
        u = H-np.log(n)
        v = H-np.log(n+1)
    return(u,v)
```

Après exécution de cette fonction, on obtient en entrant dans la console `gamma(0.001)` que $u \simeq 0.5777156$ et $v \simeq 0.5767161$. Donc $\gamma \simeq 0.57$. Ceci correspond bien à la valeur de la constante d'Euler donnée dans l'exercice 1 de la séance d'approfondissement 4.