



Thème : Simulations des variables aléatoires



Exercice 1. ♦ Lois de Pareto

D'après EMLyon Voie E 2020

Soient a et b deux réels strictement positifs. On définit la fonction f sur $\mathbb{R}$ par :

$$f : x \mapsto \begin{cases} 0 & \text{si } x > b \\ a \frac{b^a}{x^{a+1}} & \text{si } x \geq b \end{cases}$$

1. Montrer que f est une densité de probabilité.

On dit qu'une variable aléatoire suit la loi de Pareto de paramètres a et b lorsqu'elle admet pour densité la fonction f . Dans toute la suite de l'exercice, on considère une variable aléatoire X suivant la loi de Pareto de paramètres a et b .

2. Déterminer la fonction de répartition de X .
3.
 - a) Soit U une variable aléatoire suivant la loi uniforme sur $[0, 1[$. Montrer que la variable aléatoire $bU^{-\frac{1}{a}}$ suit la loi de Pareto de paramètres a et b .
 - b) En déduire une fonction Python d'en-tête `function X = pareto (a, b)` qui prend en arguments deux réels a et b strictement positifs et qui renvoie une simulation de la variable aléatoire X .

Editeur

```
import numpy as np

def pareto(a,b):
    u=np.random.rand()
    return b*u**(-1/a)
```

- 3.(c) On considère la fonction Python ci-dessous. Que contient la liste L renvoyée par la fonction **mystere**?

Editeur

```
def mystere(a,b):
    L=[]
    for p in range(2,7):
        S=0
        for k in range(1,10**p):
            S+=pareto(a,b)
        L.append(round(S/10**p,5))
    return L
```

On exécute la fonction précédente avec différentes valeurs de a et de b . Comment interpréter les résultats obtenus?

Editeur

```
>>> mystere(2,1)
[1.86379, 1.96188, 2.03994, 1.97986, 1.99862]
>>> mystere(3,2)
[3.01753, 3.08953, 2.98778, 2.99568, 2.99862]
>>> mystere(1,4)
[70.81249, 27.58769, 52.00457, 79.28056, 54.68078]
>>> mystere(1,4)
[29.62632, 45.92403, 50.88489, 79.99194, 58.12864]
```

- 4.(a) Montrer que X admet une espérance si et seulement si $a > 1$ et que, dans ce cas :

$$E(X) = \frac{ab}{a-1}.$$

4.(b) Montrer que X admet une variance si et seulement si $a > 2$ et que, dans ce cas :

$$V(X) = \frac{ab^2}{(a-1)^2(a-2)}.$$

Exercice 2. ♦♦

D'après la partie 3 de ESSEC 2014 E

Dans tout le sujet, $I =]a, b[$ est un intervalle ouvert non vide de $\mathbb{R}$, où a et b sont réels ou infinis. On dit qu'une densité de probabilité f vérifie l'hypothèse CSP(I) lorsque f est :

- continue sur I ,
- strictement positive sur I ,
- nulle en dehors de I .

On écrira alors simplement : f est CSP(I).

La plupart des langages informatiques possèdent un générateur de nombres aléatoires. En Python par exemple, on dispose de l'instruction `np.random.rand()` de la bibliothèque **numpy** (alias **np**). Ces générateurs produisent une suite de variables aléatoires indépendantes de loi uniforme sur $]0, 1[$.

On propose dans la suite deux méthodes permettant de simuler des lois continues quelconques en utilisant ces générateurs aléatoires.

Jusqu'à la fin du problème : on note Z une variable aléatoire continue à valeurs dans I , de fonction de répartition G et admettant une densité g qui est CSP(I).

a- Simulation par la méthode d'inversion

1. a) On note H la restriction de G à I . Montrer que H réalise une bijection de I sur $]0, 1[$. On note H^{-1} la bijection réciproque. Dresser le tableau de variations de H^{-1} .

Soit U une variable aléatoire suivant la loi uniforme sur $]0, 1[$. On pose $X = H^{-1}(U)$, et on note F la fonction de répartition de X .

- b) Montrer que pour tout x dans I , $F(x) = G(x)$.
- c) En déduire que X suit la même loi que Z .

2. *Simulation de lois exponentielles.*

On suppose dans cette question que Z suit la loi exponentielle de paramètre $\lambda > 0$.

- a) Expliciter l'intervalle I et les fonctions g , G et H^{-1} .
- b) Écrire une fonction Python d'en-tête **def expo(lambda)** : qui simule la loi exponentielle.

3. *Simulation de la loi de Laplace.*

On cherche dans cette question à simuler une variable aléatoire de densité g donnée par :

$$\forall x \in \mathbb{R}, \quad g(x) = \frac{1}{2} e^{-|x|} \quad (\text{densité de Laplace}).$$

Soit Y une variable aléatoire suivant la loi exponentielle de paramètre 1.

Soit V une variable aléatoire indépendante de Y suivant la loi uniforme sur $\{-1, 1\}$, ce qui signifie $\mathbf{P}(V = -1) = \mathbf{P}(V = 1) = \frac{1}{2}$. On pose $X = VY$.

- a) Vérifier que g est une densité de probabilité qui est CSP($\mathbb{R}$).
- b) Établir :
 - pour tout $x \geq 0$, $\mathbf{P}(X > x) = \frac{1}{2} \mathbf{P}(Y > x)$;
 - pour tout $x \leq 0$, $\mathbf{P}(X \leq x) = \frac{1}{2} \mathbf{P}(Y \geq -x)$.
- c) En déduire une expression de la fonction de répartition de X .
- d) Conclure que X est une variable aléatoire continue admettant g comme densité.
- e) Écrire une fonction Python qui simule la loi de Laplace.

b- Simulation par la méthode du rejet

Dans la méthode dite du rejet, pour simuler la loi de Z de densité g , on commence par déterminer une loi de probabilité que l'on sait simuler, de densité f qui est CSP(I), et qui vérifie : il existe une constante $c > 0$ telle que $\forall x \in I, g(x) \leq cf(x)$.

4. Montrer qu'il existe une fonction h continue sur I et à valeurs dans $[0, 1]$ telle que, pour tout $x \in I, g(x) = cf(x)h(x)$. On considère alors :
- une suite de variables aléatoires $(U_k)_{k \in \mathbb{N}^*}$ qui suivent la loi uniforme sur $]0, 1[$.
 - une suite de variables aléatoires $(X_k)_{k \in \mathbb{N}^*}$ à valeurs dans $]a, b[$, ayant toutes la même loi de densité de probabilité f et de fonction de répartition F .

On suppose de plus que pour tout entier $n \geq 1$, les variables $X_1, \dots, X_n, U_1, \dots, U_n$ sont mutuellement indépendantes. On définit N la variable aléatoire prenant comme valeur le premier indice k vérifiant $U_k \leq h(X_k)$.

5. ADMIS. En utilisant la partie 1, prouver l'égalité, pour tout $k \in \mathbb{N}^*$, $\mathbf{P}([U_k \leq h(X_k)]) = \frac{1}{c}$.
En déduire que N suit une loi géométrique dont on précisera le paramètre, l'espérance et la variance.
On définit la variable aléatoire X comme étant la valeur de X_N , c'est-à-dire la valeur de X_k pour le premier indice k vérifiant $U_k \leq h(X_k)$.
6. Soit $x \in I$.
- a) Soit $n \in \mathbb{N}^*$. Exprimer l'événement $[X \leq x] \cap [N = n]$ à partir des événements $[X_n \leq x] \cap [U_n \leq h(X_n)]$ et $[U_k > h(X_k)]$ pour $k \in [1, n-1]$.
 - b) (ADMIS). En utilisant la question 3.(b), montrer que pour tout $n \in \mathbb{N}^*$:

$$\mathbf{P}([X_n \leq x] \cap [U_n \leq h(X_n)]) = \frac{1}{c} G(x).$$

- c) En déduire $\mathbf{P}([X \leq x] \cap [N = n])$ en fonction de c et de $G(x)$.
- d) Montrer finalement : $\mathbf{P}([X \leq x]) = G(x)$.

7. Conclure.

8. Simulation de la loi normale.

Dans cette question, Z suit la loi normale centrée et réduite, donc $I = \mathbb{R}$.

Soit f la densité de Laplace (question 3), définie par $\forall x \in \mathbb{R}, f(x) = \frac{1}{2}e^{-|x|}$.

- a) Donner une densité g de Z qui est CSP($\mathbb{R}$).
- b) Étudier les variations sur $[0, +\infty[$ de la fonction $a : x \mapsto e^{x - \frac{x^2}{2}}$.
- c) Expliciter une constante $c > 0$ telle que, pour tout $x \geq 0 : g(x) \leq \frac{c}{2}e^{-x}$.
- d) En déduire que pour tout x réel, $g(x) \leq cf(x)$.
- e) Expliquer alors comment mettre en place la méthode du rejet pour simuler la loi normale centrée et réduite. On explicitera en particulier la fonction h introduite à la question 4.
- f) Compléter la fonction Python suivante pour qu'elle simule la loi normale centrée réduite :

```
def simulnormaleCR():
    ...

    while ... :

        x=laplace()
        u=np.random.rand()
        ...
    return normale=
```

- g) Comment simuler une loi normale $\mathcal{N}(\mu, \sigma^2)$?