

Devoir maison n°4bis
À rendre pour le 20/02

Exercice 1

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 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).

Dans le cas où a (resp. b) est fini, on admettra que l'intégrale $\int_a^b$ (resp. $\int^b$) peut se calculer comme si la fonction était continue sur $[a, b[$ (resp. $]a, b]$).

On admettra que les principaux résultats du cours concernant l'indépendance des variables aléatoires discrètes s'appliquent également aux variables aléatoires à densité.

Partie 1 - Calcul d'une probabilité

On considère dans cette partie :

- X une variable aléatoire réelle à densité, à valeurs dans I , de fonction de répartition F et admettant une densité de probabilité f qui est CSP(I).
- U une variable aléatoire qui suit une loi uniforme sur $]0, 1[$ et qui est indépendante de X .
- h une fonction continue sur I à valeurs dans $[0, 1]$.

On se propose d'établir la formule suivante :

$$\mathbb{P}([U \leq h(X)]) = \mathbb{P}(U < h(X)) = \int_a^b f(t)h(t) dt$$

On définit sur I la fonction Ψ par : $\Psi(x) = \mathbb{P}([X \leq x] \cap [U \leq h(X)])$

1. Pour tous réels x et y dans I tels que $x < y$, on pose $M(x, y) = \max_{t \in [x, y]} h(t)$ et $m(x, y) = \min_{t \in [x, y]} h(t)$.

(a) Soit x dans I . Justifier que pour tout y dans l'intervalle $]x, b[$, il existe α_y dans l'intervalle $[x, y]$ tel que $M(x, y) = h(\alpha_y)$.

(b) En déduire : $\lim_{\substack{y \rightarrow x \\ y > x}} M(x, y) = h(x)$.

(c) Montrer de même que, pour tout y dans I , $\lim_{\substack{x \rightarrow y \\ x < y}} M(x, y) = h(y)$.

On montrerait de manière analogue (on ne demande pas de le vérifier) : $\lim_{\substack{y \rightarrow x \\ y > x}} m(x, y) = h(x)$ et

$\lim_{\substack{x \rightarrow y \\ x < y}} m(x, y) = h(y)$.

2. Soit x et y deux réels de I tels que $x < y$.

- (a) Établir l'inclusion suivante entre événements :

$$([x < X \leq y] \cap [U \leq h(X)]) \subset ([x < X \leq y] \cap [U \leq M(x, y)])$$

En déduire l'inégalité :

$$\Psi(y) - \Psi(x) \leq (F(y) - F(x))M(x, y)$$

- (b) Établir une minoration analogue pour $\Psi(y) - \Psi(x)$, puis l'encadrement

$$\frac{F(y) - F(x)}{y - x} m(x, y) \leq \frac{\Psi(y) - \Psi(x)}{y - x} \leq \frac{F(y) - F(x)}{y - x} M(x, y)$$

- (c) Montrer que Ψ est dérivable sur I , et exprimer sa dérivée en fonction de f et h .

3. (a) En déduire que, pour tout x et y dans I :

$$\Psi(y) - \Psi(x) = \int_x^y f(t)h(t) dt$$

- (b) Établir : pour tout x dans I , $\Psi(x) \leq F(x)$, puis montrer : $\lim_{x \rightarrow a} \Psi(x) = 0$. En déduire :

$$\forall x \in I, \Psi(x) = \int_a^x f(t)h(t) dt$$

- (c) Établir : pour tout x dans I , $\mathbb{P}([X > x] \cap [U \leq h(X)]) = \mathbb{P}([U \leq h(X)]) - \Psi(x)$.

En déduire : $\lim_{x \rightarrow b} \Psi(x) = \mathbb{P}([U \leq h(X)])$ puis

$$\mathbb{P}([U \leq h(X)]) = \int_a^b f(t)h(t) dt$$

4. (a) Montrer que $1 - U \hookrightarrow \mathcal{U}(]0, 1[)$.

- (b) Montrer que $\mathbb{P}([U < h(X)]) = 1 - \mathbb{P}([1 - U \leq 1 - h(X)])$, en déduire :

$$\mathbb{P}(U < h(X)) = \int_a^b f(t)h(t) dt$$

Partie 2 - Simulation de variables aléatoires

La plupart des langages informatiques possèdent un générateur de nombres aléatoires. En Python par exemple, on dispose de l'instruction `rd.random()`. 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 :

- $I =]a, b[$ (a et b finis ou infinis)
- Z une variable aléatoire à valeurs dans I , de fonction de répartition G et admettant une densité g qui est CSP(I).

2a – Simulation par la méthode d'inversion

5. (a) On note H la restriction de G à $]a, b[$. Montrer que H réalise une bijection de $]a, b[$ sur $]0, 1[$.

On note H^{-1} la bijection réciproque. Dresser le tableau de variation 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 $]a, b[$, $F(x) = G(x)$.

(c) En déduire que X suit la même loi que Z .

6. Simulation de lois exponentielles.

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

(a) Expliciter l'intervalle I et les fonctions g , G et H^{-1} .

(b) En n'utilisant que le générateur aléatoire `rd.random()`, écrire une fonction Python `expo(1a)` qui simule la loi exponentielle de paramètre $1a$.

7. Simulation de la loi de Laplace.

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

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

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

Soit V une variable aléatoire indépendante de Y suivant la loi uniforme sur $\{-1, 1\}$, ce qui signifie :

$$\mathbb{P}(V = -1) = \mathbb{P}(V = 1) = \frac{1}{2}$$

On pose $X = VY$.

(a) Vérifier que g est une densité de probabilité qui est CSP($\mathbb{R}$).

(b) Établir :

- pour tout $x \geq 0$, $\mathbb{P}(X > x) = \frac{1}{2} \mathbb{P}(Y > x)$;
- pour tout $x \leq 0$, $\mathbb{P}(X \leq x) = \frac{1}{2} \mathbb{P}(Y \geq -x)$;

(c) En déduire une expression de la fonction de répartition de X .

(d) Conclure que X est une variable aléatoire admettant g comme densité.

(e) Compléter la fonction Python suivante pour qu'elle simule la loi de Laplace.

On rappelle que, pour tout réel $A > 0$, l'instruction `rd.exponential(A)` renvoie un tirage d'une variable aléatoire suivant une loi exponentielle d'espérance A .

```
def laplace():
    Y = .....
    if rd.random() .... :
        X = Y
    else:
        ....
    return X
```

2b - Simulation par la méthode du rejet

Dans la méthode dite du rejet, pour simuler la loi de Z de densité g (voir les notations en préambule de la partie 2), on commence par déterminer une loi de probabilité que l'on sait simuler, de densité f qui est CSP($]a, b[$), et qui vérifie : il existe une constante $c > 0$ telle que : $\forall x \in I, g(x) \leq cf(x)$.

8. Montrer qu'il existe une fonction h continue sur $]a, b[$ et à valeurs dans $[0, 1]$ telle que, pour tout $x \in]a, b[$, $g(x) = cf(x)h(x)$.

On considère alors :

- une suite de variable aléatoires $(U_k)_{k \in \mathbb{N}^*}$ qui suivent une loi uniforme sur $]0, 1[$.
- une suite de variable 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 $U_k \leq h(X_k)$.

9. En utilisant la partie 1, prouver l'égalité, pour tout $k \in \mathbb{N}^*$: $\mathbb{P}([U_k \leq h(X_k)]) = \frac{1}{c}$.

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 $U_k \leq h(X_k)$.

10. Soit $x \in]a, b[$.

- (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 $k \in \llbracket 1, n-1 \rrbracket$.

- (b) En utilisant la question 3b, montrer que, pour tout $n \in \mathbb{N}^*$:

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

(on rappelle que G est la fonction de répartition de Z).

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

- (d) Montrer finalement : $\mathbb{P}([X \leq x]) = G(x)$.

11. Conclure que $F_X = F_Z$.

12. Simulation de la loi normale.

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

Soit f la densité de Laplace (question 7), 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 $a : x \mapsto e^{x-x^2/2}$.

- (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, $g(x) \leq c f(x)$.

- (e) Expliquer alors comment mettre en place la méthode du rejet pour simuler la loi normale centrée réduite. On explicitera la fonction h introduite à la question 8.

- (f) Compléter la fonction Python suivante pour qu'elle simule la loi normale centrée réduite :

```
def normale():
    X = .....
    U = .....
    while ..... :
        X = .....
        U = .....
    return X
```