

Devoir maison n°4bis Corrigé

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 $M(x, y) = h(\alpha_y)$.

h est continue sur I , donc pour tout $y \in]x, b[$, h est continue sur le segment $[x, y]$; elle est donc bornée sur ce segment et atteint ses bornes. Notamment, il existe α_y dans l'intervalle $[x, y]$ tel que $h(\alpha_y) = \max_{t \in [x, y]} h(t) = M(x, y)$.

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

Faisons tendre y vers x par valeurs supérieures. Comme $x \leq \alpha_y \leq y$, on a alors $\alpha_y \rightarrow x$ par gendarmes, et donc $M(x, y) = h(\alpha_y) \rightarrow h(x)$ par continuité de h en x .

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

De même, si y tend vers x par valeurs inférieures, on a $\alpha_y \rightarrow y$, et $M(x, y) = h(\alpha_y) \rightarrow h(y)$ par continuité de h en 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 évènements :**

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

En déduire l'inégalité :

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

Supposons que $x < X \leq y$, et que $U \leq h(X)$; on a alors évidemment $x < X \leq y$, et comme $X \in]x, y] \subset [x, y]$ on a $h(X) \leq \sup_{[x, y]} h(t) = M(x, y)$; et par suite $U \leq M(x, y)$. On a donc bien l'inclusion entre évènements demandée.

Cette inclusion implique l'inégalité entre probabilités :

$$\mathbb{P}([x < X \leq y] \cap [U \leq h(X)]) \leq \mathbb{P}([x < X \leq y] \cap [U \leq M(x, y)])$$

Par indépendance de X et U le terme de droite devient :

$$\mathbb{P}([x < X \leq y] \cap [U \leq M(x, y)]) = \mathbb{P}([x < X \leq y]) \times \mathbb{P}([U \leq M(x, y)])$$

Par définition de F , on a : $\mathbb{P}([x < X \leq y]) = F(y) - F(x)$; et comme h est à valeurs dans $[0, 1]$ on a $M(x, y) \in [0, 1]$; et comme $U \hookrightarrow \mathcal{U}([0, 1])$ on peut conclure que $\mathbb{P}([U \leq M(x, y)]) = M(x, y)$.

Ainsi :

$$\mathbb{P}([x < X \leq y] \cap [U \leq M(x, y)]) = (F(y) - F(x))M(x, y)$$

Pour le terme de gauche *on ne peut pas utiliser l'indépendance* car l'évènement $(U \leq h(X))$ dépend de U et de X !!

Il faut raisonner au niveau des évènements : on a l'union disjointe :

$$[X \leq y] \cap [U \leq h(X)] = [X \leq x] \cap [U \leq h(X)] \sqcup [x < X \leq y] \cap [U \leq h(X)]$$

ce qui donne

$$\mathbb{P}([X \leq y] \cap [U \leq h(X)]) = \mathbb{P}([X \leq x] \cap [U \leq h(X)]) + \mathbb{P}([x < X \leq y] \cap [U \leq h(X)])$$

et donc

$$\Psi(y) = \Psi(x) + \mathbb{P}([x < X \leq y] \cap [U \leq h(X)])$$

et le terme de gauche vaut bien $\Psi(y) - \Psi(x)$.

On a donc bien

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

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

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

On montre cette fois l'inclusion

$$[x < X \leq y] \cap [U \leq h(X)] \supset [x < X \leq y] \cap [U \leq m(x, y)]$$

C'est essentiellement le même raisonnement ; avec cette fois que si $x < X \leq y$, on a $h(X) \geq m(x, y)$ et donc $U \leq m(x, y) \Rightarrow U \leq m(x, y) \leq h(X)$.

Au niveau des probas :

$$\mathbb{P}([x < X \leq y] \cap [U \leq h(X)]) \geq \mathbb{P}([x < X \leq y] \cap [U \leq m(x, y)])$$

et par les mêmes opérations ceci devient :

$$\Psi(y) - \Psi(x) \geq (F(y) - F(x))m(x, y)$$

En divisant par $y - x > 0$, on obtient bien l'encadrement demandé.

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

Pour $y \rightarrow x$ par valeurs supérieures : $m(x, y) \rightarrow h(x)$, $M(x, y) \rightarrow h(x)$; et $\frac{F(y) - F(x)}{y - x} \rightarrow F'(x) = f(x)$ car f est continue en x .

Dans l'encadrement précédent, cette limite fournit par théorème des gendarmes :

$$\lim_{y \rightarrow x} \frac{\Psi(y) - \Psi(x)}{y - x} = f(x)h(x)$$

et on en déduit que Ψ est dérivable en x , de nombre dérivé $f(x)h(x)$.

Ceci valant pour tout x de I , on a montré que Ψ est dérivable sur I , de dérivée $\Psi' = hf$.

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

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

C'est immédiat car Ψ est une primitive de hf d'après la question précédente.

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

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

Par définition, $\Psi(x) = \mathbb{P}([X \leq x] \cap [U \leq h(X)])$. Or on a l'inclusion entre événements :

$$[X \leq x] \cap [U \leq h(X)] \subset [X \leq x]$$

donc $\mathbb{P}([X \leq x] \cap [U \leq h(X)]) \leq \mathbb{P}([X \leq x])$ ce qui donne bien

$$\Psi(x) \leq F(x)$$

$\Psi(x)$ étant une probabilité, on a clairement $\Psi(x) \geq 0$; on peut alors essayer d'utiliser un théorème des gendarmes en utilisant l'inégalité qui vient d'être démontrée.

Pour tout x de $\mathbb{R}$, on a $F(x) = \int_{-\infty}^x f(t) dt = \int_a^x f(t) dt$ car f est nulle hors de $]a, b[$.

Cette dernière intégrale est convergente en a ; on a alors $\lim_{x \rightarrow a} \int_a^x f(t) dt = 0$; d'où $\lim_{x \rightarrow a} F(x) = 0$. Le théorème des gendarmes fournit alors

$$\lim_{x \rightarrow a} \Psi(x) = 0$$

(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

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

On utilise les probas totales avec le SCE $([X \leq x], [X > x])$:

$$\mathbb{P}([U \leq h(X)]) = \mathbb{P}([X \leq x] \cap [U \leq h(X)]) + \mathbb{P}([X > x] \cap [U \leq h(X)]) = \Psi(x) + \mathbb{P}([X > x] \cap [U \leq h(X)])$$

ce qui donne bien $\mathbb{P}([X > x] \cap [U \leq h(X)]) = \mathbb{P}([U \leq h(X)]) - \Psi(x)$.

Pour $x \rightarrow b$: $\mathbb{P}([X > x] \cap [U \leq h(X)]) \leq \mathbb{P}(X > x) = \int_x^{+\infty} f(t) dt = \int_x^b f(t) dt \rightarrow 0$, et en mino-
rant par 0, on a bien

$$\lim_{x \rightarrow b} \mathbb{P}([X > x] \cap [U \leq h(X)]) = 0$$

et donc $\lim_{x \rightarrow b} \Psi(x) = \mathbb{P}([U \leq h(X)])$.

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

$U(\Omega) =]0, 1[$ donc $(1 - U)(\Omega) =]0, 1[$ également. Donc F_{1-U} est nulle sur $\mathbb{R}_-^*$ et sur $[1, +\infty[$.
Pour $x \in [0, 1[$:

$$\begin{aligned} F_{1-U}(x) &= \mathbb{P}(1 - U \leq x) = \mathbb{P}(U \geq 1 - x) \\ &= 1 - \mathbb{P}(U < 1 - x) \\ &= 1 - \mathbb{P}(U \leq 1 - x) \quad (\text{densité}) \\ &= 1 - (1 - x) \quad (1 - x \in [0, 1]) \end{aligned}$$

$$\mathbb{P}(1 - U \leq x) = x$$

On trouve ben

$$F_{1-U}(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

ce qui montre 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 :**

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

Par passage à l'événement contraire :

$$\mathbb{P}([U < h(X)]) = 1 - \mathbb{P}([U \geq h(X)]) = 1 - \mathbb{P}([1 - U \leq 1 - h(X)])$$

Posons alors :

- $V = 1 - U$: on vient de voir que $V \hookrightarrow \mathcal{U}([0, 1])$; de plus V et X sont indépendantes par lemme des coalitions.
- $h_1(x) = 1 - h(x)$: h_1 est continue sur I , à valeurs dans $[0, 1]$.

Ce qui précède s'applique alors à V et h_1 : on trouve

$$\begin{aligned} \mathbb{P}([1 - U \leq h_1(X)]) &= \int_a^b f(t)h_1(t) dt \\ &= \int_a^b f(t)(1 - h(t)) dt \\ &= \int_a^b f(t) dt - \int_a^b f(t)h(t) dt \\ &= 1 - \int_a^b f(t)h(t) dt \end{aligned}$$

et on trouve

$$\mathbb{P}(U < h(X)) = 1 - \mathbb{P}(1 - U \leq h_1(x)) = \int_a^b f(t)h(t)dt$$

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} .

g est nulle en dehors de $I =]a, b[$, donc

$$\forall x < a, G(x) = 0 \quad \text{et} \quad \forall x > b, G(x) = 1$$

G étant continue on a aussi $\lim_{x \rightarrow a^+} G(x) = 0$ et $\lim_{x \rightarrow b^-} G(x) = 1$.

G est continue, strictement croissante sur $]a, b[$ (car g y est strictement positive) ; donc elle réalise une bijection de $]a, b[$ sur $]0, 1[$; ce qui se formule aussi en disant que la restriction de G à $]a, b[$ est une bijection de $]a, b[$ sur $]0, 1[$.

H^{-1} est alors définie sur $]0, 1[$, à valeurs dans $]a, b[$, et de même sens de variation que H ; donc strictement croissante. On obtient le tableau :

x	a	b
$H^{-1}(x)$	0	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 $]a, b[$, $F(x) = G(x)$.

$\forall x \in I, F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(H^{-1}(U) \leq x)$, avec $U \in]0, 1[$ et $x \in I$; ce qui permet d'écrire (par croissance de la bijection H) : $\forall x \in I, F(x) = \mathbb{P}(H^{-1}(U) \leq x) = \mathbb{P}(U \leq H(x))$.

Enfin, $\forall x \in \mathbb{R}, H(x) \in]0, 1[$ donc $\mathbb{P}(U \leq H(x)) = H(x)$ (loi uniforme). On a donc montré :

$$\forall x \in I, F(x) = H(x) = G(x)$$

car H et G coïncident sur I .

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

Il faut montrer que $F(x) = G(x)$ pour $x \notin I$.

Si $x \leq a$, $G(x) = 0$ (car Z est à valeurs dans I) ; et $F(x) = \mathbb{P}(H^{-1}(U) \leq x) \leq \mathbb{P}(H^{-1}(U) \leq a) = 0$ car H^{-1} est à valeurs dans $]a, b[$. Donc $F(x) = G(x) = 0$.

De même, si $x \geq b$, $G(x) = 1$; et $F(x) = \mathbb{P}(H^{-1}(U) \leq x) = 1$ car H^{-1} est à valeurs dans $]a, b[$. Donc $F(x) = G(x) = 1$.

(NB : si $a = -\infty$ et/ou $b = +\infty$, les cas ci-dessus n'ont pas lieu... mais on n'en a pas besoin !) Finalement, on a montré que $F(x) = G(x)$ pour tout réel x : X et Z ont même fonction de répartition, donc même loi.

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} .**

La densité usuelle de la loi exponentielle est nulle sur $\mathbb{R}_-$ et > 0 sur $\mathbb{R}_+^*$: on a donc $I = \mathbb{R}_+^*$; $g(x) = 0$ pour $x \leq 0$ et $g(x) = \lambda e^{-\lambda x}$ pour $x > 0$; $G(x) = 0$ pour $x \leq 0$ et $G(x) = 1 - e^{-\lambda x}$ pour $x > 0$.

H est donc la fonction définie de $\mathbb{R}_+^*$ dans $]0, 1[$ par $H(x) = 1 - e^{-\lambda x}$; on voit alors que $y = H(x) \Leftrightarrow x = -\frac{1}{\lambda} \ln(1 - y)$; de sorte que $H^{-1}(x) = -\frac{1}{\lambda} \ln(1 - x)$.

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

Il suffit de considérer une variable U suivant $\mathcal{U}([0, 1])$, et de renvoyer $H^{-1}(U)$:

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

NB : vous reconnaissez évidemment dans ce cas particulier une question très très usuelle !!

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($\mathbb{R}$).**

g est une densité car elle est continue (composée de fonctions continues sur $\mathbb{R}$) et positive (> 0 car c'est une exponentielle). Montrons que $\int_{-\infty}^{+\infty} g(x) dx$ converge et vaut 1.

On remarque que g est paire : il suffit donc de s'intéresser à $\int_0^{+\infty} g(x) dx = \frac{1}{2} \int_0^{+\infty} e^{-x} dx$ qui converge (intégrale usuelle) et vaut $\frac{1}{2}$. Par parité, $\int_{-\infty}^{+\infty} g(x) dx$ converge et vaut 1 : g est bien une densité.

Pour le caractère CSP($\mathbb{R}$), il faut montrer que g est continue et > 0 sur $\mathbb{R}$, ce qui a été fait.

(b) **Établir :**

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

On utilise une formule des probabilités totales, avec le SCE $([V = 1], [V = -1])$. On a, pour $x \geq 0$:

$$\begin{aligned} \mathbb{P}(X > x) &= \mathbb{P}(X > x, V = 1) + \mathbb{P}(X > x, V = -1) \\ &= \mathbb{P}(VY > x, V = 1) + \mathbb{P}(VY > x, V = -1) \\ &= \mathbb{P}(Y > x, V = 1) + \mathbb{P}(-Y > x, V = -1) \\ &= \mathbb{P}(Y > x) \mathbb{P}(V = 1) + \mathbb{P}(Y < -x) \mathbb{P}(V = -1) \quad \text{par indépendance de } V \text{ et } Y \\ &= \frac{1}{2} \mathbb{P}(Y > x) + \frac{1}{2} \mathbb{P}(Y < -x) \end{aligned}$$

Or $Y \mapsto \mathcal{E}(1)$ donc est à valeurs positives ; pour $x \geq 0$ on a $\mathbb{P}(Y < -x) = 0$; et donc $\mathbb{P}(X > x) = \frac{1}{2} \mathbb{P}(Y > x)$.

De manière similaire, pour tout $x \leq 0$,

$$\mathbb{P}(X \leq x) = \mathbb{P}(X \leq x, V = 1) + \mathbb{P}(X \leq x, V = -1) = \frac{1}{2} \mathbb{P}(Y \leq x) + \frac{1}{2} \mathbb{P}(Y \geq -x) = \frac{1}{2} \mathbb{P}(Y \geq -x)$$

car comme Y est à valeurs positives et $x \leq 0$, on a $\mathbb{P}(Y \leq x) = 0$.

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

Pour tout $x \leq 0$, on obtient directement

$$F_X(x) = \mathbb{P}([X \leq x]) = \frac{1}{2} \mathbb{P}([Y \geq -x]) = \frac{1}{2} (1 - F_Y(-x))$$

Pour $x > 0$,

$$F_X(x) = \mathbb{P}([X \leq x]) = 1 - \mathbb{P}([X > x]) = 1 - \frac{1}{2} \mathbb{P}([Y > x]) = 1 - \frac{1}{2} (1 - F_Y(x)) = \frac{1}{2} (1 + F_Y(x))$$

et avec l'expression de F_Y comme fonction de répartition de la loi $\mathcal{E}(1)$:

$$\begin{aligned} \forall x \leq 0, F_X(x) &= \frac{1}{2} (1 - (1 - e^{-(-x)})) = \frac{1}{2} e^x \\ \forall x > 0, F_X(x) &= \frac{1}{2} (1 + 1 - e^{-x}) = 1 - \frac{1}{2} e^{-x} \end{aligned}$$

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

La fonction de répartition est $\mathcal{C}^1$ sur $\mathbb{R}^*$, et continue en 0 (vérification aisée) ; X est donc à densité, et on obtient une densité en dérivant :

$$\forall x > 0, f_X(x) = \frac{1}{2} e^{-x} \quad ; \quad \forall x < 0, f_X(x) = \frac{1}{2} e^x$$

et en posant $f_X(0) = \frac{1}{2}$ on a bien $f_X = g$.

(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 .

Il suffit donc de tirer $Y \hookrightarrow \mathcal{E}(1)$, de tirer V valant ± 1 de manière équiprobable, et de renvoyer VY ; ou encore de tirer Y , et de renvoyer Y ou $-Y$ de manière équiprobable.

```
def laplace():
    Y = rd.exponential(1)
    if rd.random() < 1/2 :
        X = Y
    else:
        X = -Y
    return X
```

NB : en fait

```
def laplace():
    Y = rd.exponential(1)
    if rd.random() < 1/2 :
        return Y
    return -Y
```

eût été bien plus esthétique...

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 : $\forall x \in I, g(x) \leq c f(x)$.

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

Il semble raisonnable de poser $h(x) = \frac{g(x)}{c f(x)}$. Montrons que cette fonction convient :

- f est CSP(I) donc > 0 sur I , et $c > 0$. La fonction h ainsi définie est donc continue sur I comme composée de fonctions continues dont le dénominateur ne s'annule pas.
- en divisant car $c f(x) > 0$ l'inégalité supposée dans cette question, on obtient :

$$\forall x \in I, \underbrace{\frac{g(x)}{c f(x)}}_{=h(x)} \leq 1$$

De plus, f et g sont des densités donc sont positives ; $c > 0$; par suite h est bien à valeurs positives.

- enfin, par construction, on a bien $g(x) = c f(x) h(x)$.

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.

Pour tout k , U_k et X_k sont indépendantes. f est CSP(I), h est continue de I dans $[0, 1]$, donc la partie 1 s'applique :

$$\forall k \in \mathbb{N}^*, \mathbb{P}(U_k \leq h(X_k)) = \int_a^b f(t)h(t) dt = \int_a^b \frac{g(t)}{c} dt = \frac{1}{c} \int_a^b g(t) dt$$

Or, g est une densité CSP($]a, b[$), donc $\int_a^b g(t) dt = 1$; et on a bien $\mathbb{P}([U_k \leq h(X_k)]) = \frac{1}{c}$.

N mesure alors le premier rang k pour lequel $U_k \leq h(X_k)$; événement qui arrive avec une proba $\frac{1}{c}$; et les tirages successifs de (U_k, X_k) sont indépendants. On est donc dans un schéma de loi géométrique : $N \hookrightarrow \mathcal{G}\left(\frac{1}{c}\right)$.

Le cours donne alors $E(N) = \frac{1}{\frac{1}{c}} = c$ et $V(N) = \frac{1 - \frac{1}{c}}{\frac{1}{c^2}} = c(c-1)$.

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

N est le premier indice i tel que $U_i \leq h(X_i)$; on a donc

$$[N = n] = \left(\bigcap_{i=1}^{n-1} [U_i > h(X_i)] \right) \cap [U_n \leq h(X_n)]$$

La variable X vaut alors X_n ; d'où

$$[X \leq x] \cap [N = n] = \left(\bigcap_{i=1}^{n-1} [U_i > h(X_i)] \right) \cap [U_n \leq h(X_n)] \cap [X_n \leq x]$$

- (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).

U_n et X_n sont indépendantes, et les hypothèses nous permettent encore d'appliquer la partie 1 :

$$\begin{aligned} \mathbb{P}([X_n \leq x] \cap [U_n \leq h(X_n)]) &= \int_a^x f(t)h(t) dt \\ &= \frac{1}{c} \int_a^x g(t) dt \\ &= \frac{1}{c} (G(x) - G(a)) \end{aligned}$$

Or Z est à densité nulle hors de $[a, b]$, donc $G(a) = \mathbb{P}(Z \leq a) = 0$; et donc :

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

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

Par indépendance mutuelle des U_k, X_k :

$$\begin{aligned}\mathbb{P}([X \leq x] \cap [N = n]) &= \mathbb{P}\left(\left(\bigcap_{i=1}^{n-1} [U_i > h(X_i)]\right) \cap [U_n \leq h(X_n)] \cap [X_n \leq x]\right) \\ &= \left(\prod_{i=1}^{n-1} \mathbb{P}([U_i > h(X_i)])\right) \times \mathbb{P}([U_n \leq h(X_n)] \cap [X_n \leq x]) \\ &= \left(1 - \frac{1}{c}\right)^{n-1} \frac{1}{c} G(x) \quad \text{d'après 14 et 15b}\end{aligned}$$

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

On applique enfin la formule des probabilités totales avec le SCE $(N = n)_{n \geq 1}$:

$$\begin{aligned}\mathbb{P}(X \leq x) &= \sum_{n=1}^{+\infty} \mathbb{P}([X \leq x] \cap [N = n]) \\ &= \frac{1}{c} G(x) \sum_{n=1}^{+\infty} \left(1 - \frac{1}{c}\right)^{n-1} \\ &= \frac{1}{c} G(x) \frac{1}{1 - (1 - \frac{1}{c})} = G(x)\end{aligned}$$

11. **Conclure que** $F_X = F_Z$.

X étant défini comme l'un des X_i , on a $X(\Omega)$ égal à $]a, b[$ (image commune de tous les X_i).

$X(\Omega) = Z(\Omega) =]a, b[$, donc les fonctions de répartition de Z et X sont nulles sur $] -\infty, a]$ et sur $[b, +\infty[$.

Sur $]a, b[$, on a vu que $\mathbb{P}(X \leq x) = G(x)$ avec G fonction de répartition de Z : ici aussi les fonctions de répartition de Z et X coïncident.

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 $\forall x \in \mathbb{R} \quad f(x) = \frac{1}{2} e^{-|x|}$.

(a) **Donner une densité g de Z qui est CSP($\mathbb{R}$).**

D'après le cours, une densité de Z est donnée par :

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Cette fonction est bien CSP($\mathbb{R}$) (facile à vérifier).

(b) **Étudier les variations sur $]0, +\infty[$ de la fonction $a : x \mapsto e^{x-x^2/2}$.**

Cette fonction est dérivable, et $a'(x) = (1 - x)e^{x-x^2/2}$. a est donc croissante sur $]0, 1]$ et décroissante sur $[1, +\infty[$; elle atteint son max en 1, et ce max vaut $e^{1/2}$.

- (c) **Expliciter une constante $c > 0$ telle que, pour tout $x \geq 0$: $g(x) \leq \frac{c}{2} e^{-x}$.**

On vient donc de montrer, pour tout $x \geq 0$:

$$\begin{aligned} a(x) &\leq e^{1/2} \\ e^{x-x^2/2} &\leq e^{1/2} \\ e^{-x^2/2} &\leq e^{1/2-x} \\ g(x) &\leq \frac{1}{\sqrt{2\pi}} e^{1/2} e^{-x} \end{aligned}$$

et $c = \frac{2}{\sqrt{2\pi}} e^{1/2}$ convient.

(NB : numériquement on a bien $c > 1$, ce qui rend pertinents tous les calculs précédents...)

- (d) **En déduire que pour tout x réel, $g(x) \leq c f(x)$.**

On a le résultat pour $x \geq 0$. Si $x < 0$, $-x > 0$ donc $g(-x) \leq c f(-x)$; or f et g sont paires donc ceci s'écrit aussi $g(x) \leq c f(x)$, et le résultat est encore vrai.

Finalement, 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.**

La fonction h est donc donnée par :

$$\forall x \in \mathbb{R}, h(x) = \frac{g(x)}{c f(x)} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{\sqrt{2\pi}}{2} e^{-1/2} 2 e^{-|x|} = \exp\left(-\left(|x| + \frac{x^2}{2} + \frac{1}{2}\right)\right)$$

Un tirage de X défini comme précédemment permet ainsi de simuler Z . Ce tirage se fait en tirant des X_i (ce qu'on sait faire par hypothèse) ; des U_i (loi uniforme : on sait faire), jusqu'à ce qu'on ait $U_i \leq h(X_i)$. Ce X_i donne alors une simulation de Z .

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

```
def normale():
    X = laplace()
    U = rd.random()
    while U > np.exp(-np.abs(X)-X**2/2-1/2):
        X = laplace()
        U = rd.random()
    return X
```