

TD 1 : SIMULATION MONTE-CARLO

EXERCICE 1 (Méthode d'inversion)

1. Soit $\lambda > 0$. Rappeler la définition d'une fonction mesurable $\varphi_\lambda :]0, 1[\rightarrow \mathbb{R}$, telle que si $U \sim \text{Unif}([0, 1])$, alors $\varphi_\lambda(U) \sim \text{Exp}(\lambda)$.

Correction Puisque la fonction de répartition de la loi $\text{Exp}(\lambda)$ est $F : x \in \mathbb{R} \mapsto (1 - e^{-\lambda x})\mathbf{1}_{x \geq 0}$, dont l'inverse généralisée est

$$F^{-1} : u \in [0, 1] \mapsto \begin{cases} -\infty & \text{pour } u = 0 \\ -\log(1 - u)/\lambda & \text{lorsque } u \in]0, 1[\\ \infty & \text{pour } u = 1, \end{cases}$$

et que $1 - U \sim \text{Unif}([0, 1])$, il vient $\text{Exp}(\lambda) = \text{Loi}(F^{-1}(U)) = \text{Loi}(F^{-1}(1 - U)) = \text{Loi}(-\log(U)/\lambda)$. Il suffit donc de prendre, pour tout $u \in]0, 1[$, $\varphi_\lambda(u) = -\log(u)/\lambda$.

2. Donner un générateur de la loi de Pareto de paramètres $(x_m, k) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$, dont une densité est donnée par :

$$f : x \in \mathbb{R} \mapsto k \frac{x_m^k}{x^{k+1}} \mathbf{1}_{x \geq x_m}.$$

Correction On obtient par intégration que la fonction de répartition de la loi de Pareto considérée est $F : x \in \mathbb{R} \mapsto (1 - (x_m/x)^k)\mathbf{1}_{x \geq x_m}$, dont l'inverse généralisée est

$$F^{-1} : u \in [0, 1] \mapsto \begin{cases} -\infty & \text{pour } u = 0 \\ x_m/(1 - u)^{1/k} & \text{lorsque } u \in]0, 1[\\ \infty & \text{pour } u = 1. \end{cases}$$

En prenant $U \sim \text{Unif}([0, 1])$, on obtient simplement comme générateur $F^{-1}(1 - U) = x_m/U^{1/k}$.

3. Proposer une manière de simuler la loi de Cauchy de paramètres $(m, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$, dont une densité est donnée par :

$$f : x \in \mathbb{R} \mapsto \frac{\sigma}{\pi [(x - m)^2 + \sigma^2]}.$$

Correction On obtient la fonction de répartition de la loi de Cauchy considérée par intégration (en utilisant le changement de variable $y = (t - m)/\sigma$) :

$$\forall x \in \mathbb{R} : F(x) = \int_{-\infty}^x \frac{1/\sigma}{\pi \left[1 + \left(\frac{x-m}{\sigma}\right)^2\right]} dt = \int_{-\infty}^{\frac{x-m}{\sigma}} \frac{1}{\pi [1 + y^2]} dy = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x - m}{\sigma}\right).$$

On obtient alors pour inverse généralisée :

$$F^{-1} : u \in [0, 1] \mapsto \begin{cases} -\infty & \text{pour } u = 0 \\ m + \sigma \tan\left(\pi\left(u - \frac{1}{2}\right)\right) & \text{lorsque } u \in]0, 1[\\ \infty & \text{pour } u = 1. \end{cases}$$

En prenant $U \sim \text{Unif}([0, 1])$, on obtient comme générateur $F^{-1}(U) = m + \sigma \tan(\pi(U - 1/2))$.

4. Faire de même avec la loi de Bernoulli de paramètre p , $\text{Ber}(p)$, où $p \in]0, 1[$.

Correction La fonction de répartition de la loi $\text{Ber}(p)$ et son inverse généralisée valent respectivement :

$$F : x \in \mathbb{R} \mapsto \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } 0 \leq x < 1 \\ 1 & \text{si } 1 \leq x \end{cases} \quad \text{et} \quad F^{-1} : u \in [0, 1] \mapsto \begin{cases} -\infty & \text{si } u = 0 \\ 0 & \text{si } 0 < u \leq 1 - p \\ 1 & \text{si } 1 - p < u \leq 1. \end{cases}$$

En prenant $U \sim \text{Unif}([0, 1])$, on obtient comme générateur $F^{-1}(1 - U) = \mathbf{1}_{U \leq p}$.

5. Proposer une manière concrète de simuler suivant la loi géométrique de paramètre p , $\text{Geo}_{\mathbb{N}^*}(p)$, où $p \in]0, 1[$.

Correction Pour rappel, si $X \sim \text{Geo}_{\mathbb{N}^*}(p)$, alors X a comme issues $\mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = (1 - p)^{k-1}p$. En notant, pour tout $k \in \mathbb{N}^*$, $x_k = k$ et $p_k = (1 - p)^{k-1}p$, la fonction de répartition s'écrit :

$$\forall x \in \mathbb{R}, \quad F(x) = \sum_{k \geq 1} p_k \mathbf{1}_{x_k \leq x}.$$

Soit $u \in]0, 1[$, un dessin montre que :

$$F(x) \geq u \iff x \geq x_n, \text{ pour } n \in \mathbb{N}^* \text{ tel que } \sum_{k=1}^{n-1} p_k < u \leq \sum_{k=1}^n p_k.$$

Ainsi,

$$\forall u \in [0, 1] : F^{-1}(u) = \begin{cases} -\infty & \text{si } u = 0 \\ \sum_{n \geq 1} x_n \mathbf{1}_{\sum_{k=1}^{n-1} p_k < u \leq \sum_{k=1}^n p_k} & \text{si } 0 < u < 1 \\ \infty & \text{si } u = 1. \end{cases}$$

On pourrait donc prendre $U \sim \text{Unif}([0, 1])$ faire une boucle while pour déterminer $N \in \mathbb{N}^*$ tel que $\sum_{k=1}^{N-1} p_k < U \leq \sum_{k=1}^N p_k$ et retourner N . En réalité, cette démarche est excessivement compliquée car nous avons une formule explicite pour la fonction de répartition de X :

$$\forall x \in \mathbb{R}, \quad F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \leq \lfloor x \rfloor) = 1 - \mathbb{P}(X > \lfloor x \rfloor) = 1 - (1 - p)^{\lfloor x \rfloor}.$$

On obtient donc que pour tout $u \in]0, 1[$:

$$F^{-1}(u) = \inf \left\{ x \in \mathbb{R}, 1 - (1 - p)^{\lfloor x \rfloor} \geq u \right\} = \inf \left\{ x \in \mathbb{R}, \frac{\log(1 - u)}{\log(1 - p)} \leq \lfloor x \rfloor \right\} = \left\lceil \frac{\log(1 - u)}{\log(1 - p)} \right\rceil.$$

Un générateur de $\text{Geo}_{\mathbb{N}^*}(p)$ est donc $\lceil \log(U)/\log(1 - p) \rceil$.

6. Soit $p \in]0, 1[$. Montrer que si $T \sim \text{Exp}(-\log(1-p))$, alors $\lfloor T + 1 \rfloor \sim \text{Geo}_{\mathbb{N}^*}(p)$.

Correction Soit $U \sim \text{Unif}([0, 1])$. Alors, d'après la question 1, $\text{Loi}(T) = \text{Loi}(\log(U)/\log(1-p))$ puis $\text{Loi}(\lfloor T + 1 \rfloor) = \text{Loi}(\lfloor \log(U)/\log(1-p) + 1 \rfloor)$. Or $\lfloor \log(U)/\log(1-p) + 1 \rfloor = \lceil \log(U)/\log(1-p) \rceil$ p.s. donc $\lfloor T + 1 \rfloor \sim \text{Geo}_{\mathbb{N}^*}(p)$ d'après la question précédente.

On aurait aussi pu raisonner directement : pour tout $k \in \mathbb{N}^*$,

$$\mathbb{P}(\lfloor T + 1 \rfloor = k) = \mathbb{P}(k-1 \leq T < k) = \left(1 - e^{k \log(1-p)}\right) - \left(1 - e^{(k-1) \log(1-p)}\right) = (1-p)^{k-1}p.$$

Ainsi $\lfloor T + 1 \rfloor \sim \text{Geo}_{\mathbb{N}^*}(p)$.

EXERCICE 2 (Méthode du rejet)

Soit $\lambda \in]0, 1[$.

1. Décrire la méthode du rejet pour simuler la loi $\text{Poi}(\lambda)$ à partir de la loi $\text{Geo}_{\mathbb{N}}(1-\lambda)$.

Correction Les deux lois considérées sont toutes les deux absolument continues par rapport à la mesure de comptage sur $\mathbb{N}$, avec pour densités respectives :

$$f : x \in \mathbb{N} \mapsto e^{-\lambda} \frac{\lambda^x}{x!} \quad \text{et} \quad g : x \in \mathbb{N} \mapsto (1 - (1 - \lambda))^x (1 - \lambda) = \lambda^x (1 - \lambda)$$

(la loi géométrique sur $\mathbb{N}$ est la loi du nombre d'échecs avant le premier succès). Il vient donc

$$m = \sup_{x \in \mathbb{N}} \frac{f(x)}{g(x)} = \sup_{x \in \mathbb{N}} \frac{e^{-\lambda}}{x!(1-\lambda)} = \frac{e^{-\lambda}}{(1-\lambda)},$$

puis,

$$\forall x \in \mathbb{N} : \frac{f(x)}{mg(x)} = \frac{1}{x!}.$$

Ainsi, on peut considérer deux suites indépendantes (Y_n) et (U_n) de variables aléatoires i.i.d. de loi $\text{Unif}([0, 1])$, et définir $N = \min\{n \geq 1, U_n \leq 1/(\lfloor \log(Y_n)/\log(\lambda) \rfloor)!\}$ (car $\lfloor \log(Y_n)/\log(\lambda) \rfloor \sim \text{Geo}_{\mathbb{N}}(1-\lambda)$). Il vient, par la méthode du rejet : $\lfloor \log(Y_N)/\log(\lambda) \rfloor \sim \text{Poi}(\lambda)$.

2. Comment simuler simplement $Y \sim \text{Geo}_{\mathbb{N}}(1-\lambda)$?

Correction Voir ci-dessus.

3. Quelle est la probabilité de rejet ?

Correction La probabilité de rejet p_r est la probabilité que la variable aléatoire N soit strictement supérieure à 1, or on sait que $N \sim \text{Geo}_{\mathbb{N}^*}(1/m)$ donc

$$p_r = 1 - \mathbb{P}(N = 1) = 1 - \frac{1}{m} = 1 - (1-\lambda)e^\lambda.$$

EXERCICE 3 (★) Méthode du rejet optimisée)

On veut simuler une loi normale $\mathcal{N}(0, 1)$ en utilisant comme proposition une loi de Laplace de paramètre $\lambda > 0$, c'est-à-dire de densité $g_\lambda(x) = \frac{\lambda}{2}e^{-\lambda|x|}$.

1. Déterminer la valeur de λ qui permet de minimiser la probabilité de rejet. Vérifier que cette dernière vaut environ 0,24.

Correction En notant f la densité de la loi normale $\mathcal{N}(0, 1)$, commençons par noter que pour une méthode de rejet basée sur le couple (g_λ, m_λ) vérifiant pour tout x ,

$$f(x) \leq m_\lambda g_\lambda(x),$$

la probabilité de rejet est la probabilité qu'une loi géométrique de paramètre $1/m_\lambda$ soit strictement supérieure à 1, c'est-à-dire $1 - 1/m_\lambda$. Minimiser la probabilité de rejet revient donc à minimiser m_λ . Autrement dit, on cherche le couple de paramètres (λ, m_λ) tel que $f \leq m_\lambda g_\lambda$ avec m_λ minimal. Pour ceci, notons que par symétrie,

$$\forall x \in \mathbb{R}, \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \leq m_\lambda \frac{\lambda}{2}e^{-\lambda|x|} \iff \forall x \geq 0, \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \leq m_\lambda \frac{\lambda}{2}e^{-\lambda x}$$

ceci étant encore équivalent à

$$m_\lambda \geq \sup_{x \geq 0} \sqrt{\frac{2}{\pi}} \frac{1}{\lambda} e^{\lambda x - x^2/2} = \sqrt{\frac{2}{\pi}} \frac{1}{\lambda} e^{\lambda^2/2} = \varphi(\lambda).$$

Il reste à minimiser la fonction φ sur $(0, +\infty)$, ce qui est équivalent à minimiser son logarithme, ce qui donne $\lambda = 1$. Il suffit ensuite de prendre

$$m_1 = \varphi(1) = \sqrt{\frac{2e}{\pi}} \implies 1 - \frac{1}{m_1} = 1 - \sqrt{\frac{\pi}{2e}} \approx 0.24.$$

-
2. Comment simuler une variable X suivant une loi de Laplace ?

Correction Pour appliquer cet algorithme de rejet, il faut donc commencer par simuler Y suivant une loi de Laplace de paramètre 1. Il y a plusieurs possibilités : la méthode générique consiste à appliquer la méthode d'inversion. Partant de la fonction de répartition

$$F(y) = \frac{1}{2}e^y \mathbf{1}_{y < 0} + \left(1 - \frac{1}{2}e^{-y}\right) \mathbf{1}_{y \geq 0},$$

on en déduit son inverse

$$F^{-1}(u) = \log(2u) \mathbf{1}_{0 < u < 1/2} - \log(2(1-u)) \mathbf{1}_{1/2 \leq u < 1}.$$

Il suffit donc de tirer $U \sim \text{Unif}([0, 1])$ puis de poser $Y = F^{-1}(U)$.

On peut aussi voir la loi de Laplace comme un mélange équiprobable d'une exponentielle positive et d'une exponentielle négative. Ainsi, une méthode de simulation est de commencer par tirer une Rademacher R , c'est-à-dire $R = \pm 1$ de façon équiprobable, soit

$$R = -\mathbf{1}_{U \leq 1/2} + \mathbf{1}_{U \geq 1/2} \text{ avec } U \sim \text{Unif}([0, 1]).$$

On tire par ailleurs une loi $E \sim \text{Exp}(1)$, pour obtenir $Y = R \times E$ qui suit bien une loi de Laplace de paramètre 1.

Quoi qu'il en soit, après simplifications, la fonction rapport d'acceptation est alors définie par

$$r(Y) = \frac{f(Y)}{m_{\lambda}g(Y)} = e^{-\frac{1}{2}(|Y|-1)^2}.$$

Soit enfin $V \sim \text{Unif}([0, 1])$, alors si $V \leq r(Y)$, la variable $X = Y$ suit une loi normale centrée réduite. Le résultat précédent dit qu'en moyenne, environ 3 fois sur 4, la proposition Y est acceptée.

Remarque : on peut gagner un peu de temps dans la simulation par $Y = R \times E$ puisque la loi de Laplace et la loi normale sont symétriques : on tire d'abord Y_0 suivant une loi exponentielle, on regarde si la proposition est acceptée via le même rapport que ci-dessus, c'est-à-dire

$$r(Y_0) = \frac{f(Y_0)}{m_{\lambda}g(Y_0)} = e^{-\frac{1}{2}(|Y_0|-1)^2} = e^{-\frac{1}{2}(Y_0-1)^2},$$

et, pour $V \sim \text{Unif}([0, 1])$, alors si $V \leq r(Y_0)$, on tire une Rademacher R et on pose $Y = R \times Y_0$.

EXERCICE 4 (Conditionnement et inversion)

On considère le couple (X, Y) de densité

$$f : (x, y) \in \mathbb{R}^2 \mapsto yx^{y-1}e^{-y}\mathbf{1}_{y>0}\mathbf{1}_{0<x<1}.$$

1. Quelle est la loi de Y ?

Correction La densité de Y s'obtient en marginalisant par rapport à x , ce qui donne ici

$$f(y) = \int_0^1 f(x, y)dx = e^{-y}\mathbf{1}_{y>0},$$

et Y suit donc une loi exponentielle de paramètre 1.

-
2. Soit $y > 0$. Déterminer $\text{Loi}(X|Y = y)$.

Correction La densité conditionnelle de X sachant $Y = y$ vaut donc pour tout $y > 0$ fixé,

$$f(x|y) = yx^{y-1}\mathbf{1}_{0<x<1}.$$

On reconnaît la loi $\text{Beta}(y, 1)$.

-
3. Déterminez, pour tout $x \in]0, 1[$, $\mathbb{P}(X \leq x|Y = y)$ et en déduire une méthode de simulation du couple (X, Y) .

Correction Pour tout x entre 0 et 1, on a donc :

$$\mathbb{P}(X \leq x|Y = y) = \int_0^x f(x|y)dx = x^y.$$

La variable y étant fixée, on peut inverser cette fonction de répartition conditionnelle et appliquer la méthode d'inversion. Puisque

$$u = x^y \iff x = u^{1/y},$$

on en déduit que pour simuler une réalisation du couple (X, Y) , il suffit de commencer par générer Y suivant une loi exponentielle de paramètre 1 puis générer U suivant une loi uniforme sur $[0, 1]$ et enfin poser $X = U^{1/Y}$.

EXERCICE 5 (★) Conditionnement

On considère le couple (X, Y) de densité

$$f : (x, y) \in \mathbb{R}^2 \mapsto \frac{1}{\sqrt{8\pi}} e^{-y^2 x/2} e^{-\sqrt{x}} \mathbf{1}_{x>0}.$$

1. Soit $x > 0$. Quelle est la loi $\text{Loi}(Y|X = x)$?

Correction Pour déterminer la loi de Y sachant $X = x$, on utilise le symbole “ $\propto$ ” (comme en statistique bayésienne) et on reconnaît immédiatement une gaussienne centrée :

$$f(y|x) \propto e^{-xy^2/2} \implies \text{Loi}(Y|X = x) = \mathcal{N}(0, 1/x).$$

-
2. Quelle est la loi de $\sqrt{X}$?

Correction La règle de Bayes donne :

$$f(x) = \frac{f(x, y)}{f(y|x)} = \frac{1}{2\sqrt{x}} e^{-\sqrt{x}} \mathbf{1}_{x>0},$$

d'où l'on déduit la fonction de répartition de la variable $\sqrt{X}$: pour tout $r > 0$,

$$F(r) = \mathbb{P}(\sqrt{X} \leq r) = \mathbb{P}(X \leq r^2) = \int_0^{r^2} f(x) dx = \int_0^{r^2} \frac{1}{2\sqrt{x}} e^{-\sqrt{x}} dx = 1 - e^{-r},$$

via le changement de variable $t = \sqrt{x}$. Bilan : $\sqrt{X}$ suit une loi exponentielle de paramètre 1.

-
3. En déduire une méthode de simulation du couple (X, Y) .

Correction Pour simuler le couple (X, Y) : on simule $T \sim \text{Exp}(1)$, puis on considère $X = T^2$ et on simule $Y \sim \mathcal{N}(0, 1/X)$.

EXERCICE 6 (Rejet et conditionnement)

Soit X une variable aléatoire de fonction de répartition F , supposée bijective, et de densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}$. On suppose savoir simuler la loi de X .

1. Exhiber une densité de $\text{Loi}(X|X > a)$.

Correction Puisque F est bijective, on a $\mathbb{P}(X > a) > 0$ pour tout réel a . La fonction de répartition de $\text{Loi}(X|X > a)$ s'écrit :

$$\forall x \in \mathbb{R}, \quad \mathbb{P}(X \leq x|X > a) = \frac{\mathbb{P}(a < X \leq x)}{\mathbb{P}(X > a)} = \frac{F(x) - F(a)}{1 - F(a)} \mathbf{1}_{x>a}.$$

En dérivant partout sauf en $x = a$, on obtient une densité pour $\text{Loi}(X|X > a)$: pour tout $x \in \mathbb{R}$, $f_a(x) = f(x) \mathbf{1}_{x>a} / (1 - F(a))$.

2. Comment simuler $\text{Loi}(X|X > a)$ à l'aide d'une méthode de rejet ? Que se passe-t-il lorsque a devient grand ?

Correction Si l'on prend comme loi instrumentale $\text{Loi}(X)$, on a :

$$m = \sup_{x \in \mathbb{R}} \frac{f_a(x)}{f(x)} = \sup_{x \in \mathbb{R}} \frac{\mathbf{1}_{x>a}}{1 - F(a)} = \frac{1}{1 - F(a)} = \frac{1}{\mathbb{P}(X > a)},$$

puis,

$$\forall x \in \mathbb{R} : \frac{f_a(x)}{mf(x)} = \mathbf{1}_{x>a}.$$

Ainsi, pour simuler $\text{Loi}(X|X > a)$ à l'aide d'une méthode de rejet, il suffit de simuler Y_1 suivant $\text{Loi}(X)$, puis de poser $X = Y_1$ si $Y_1 > a$, et de recommencer avec une nouvelle variable Y_2 sinon, et ainsi de suite jusqu'à obtenir $Y_N > a$.

Avec les notations du cours, il s'ensuit que la variable aléatoire N suit une loi géométrique de paramètre $1/m = \mathbb{P}(X > a)$. Le nombre moyen d'essais pour obtenir une réalisation de $\text{Loi}(X|X > a)$ est donc égal à $1/\mathbb{P}(X > a)$, quantité qui tend vers 0 quand a tend vers l'infini. Cette méthode est donc à proscrire lorsque a se situe dans la queue de la distribution de X .

Remarques :

- Il n'est en réalité pas nécessaire que $\text{Loi}(X)$ soit à densité, ni que F soit bijective, pour appliquer la méthode du rejet. En effet, puisque l'idée est de simuler $\text{Loi}(X|X > a)$ grâce à la loi auxiliaire $\text{Loi}(X)$, il suffit de remarquer que toutes les deux sont absolument continues par rapport à $\text{Loi}(X)$! Les densités respectives sont $f_a : x \in \mathbb{R} \mapsto \mathbf{1}_{x>a}/\mathbb{P}(X > a)$ et $f : x \in \mathbb{R} \mapsto 1$. Bref, la seule condition importante est que $\mathbb{P}(X > a) > 0$, sans quoi l'algorithme ne s'arrête jamais...
- C'est en fait une méthode de rejet basique : pour simuler X selon la loi de X sachant $\{X \in B\}$, on simule des réalisations suivant la loi de X jusqu'à tomber dans B .

3. Soient $U \sim \text{Unif}([0, 1])$ et T définie par

$$T = F^{-1}(F(a) + (1 - F(a))U).$$

Déterminer la fonction de répartition de T et en déduire une méthode de simulation de X conditionnellement à $X > a$. Comparer à la méthode de la question précédente.

Correction Soit U une variable de loi uniforme sur $[0, 1]$ et T définie par

$$T = F^{-1}(F(a) + (1 - F(a))U).$$

Puisque $F(a) + (1 - F(a))U \geq F(a)$, la variable T est supérieure à la valeur a . Plus précisément, sa fonction de répartition F_T vaut, pour tout $t > a$,

$$F_T(t) = \mathbb{P}(T \leq t) = \mathbb{P}(F^{-1}(F(a) + (1 - F(a))U) \leq t) = \mathbb{P}(F(a) + (1 - F(a))U \leq F(t)),$$

ce qui conduit à

$$F_T(t) = \mathbb{P}\left(U \leq \frac{F(t) - F(a)}{1 - F(a)}\right) = \frac{F(t) - F(a)}{1 - F(a)}.$$

D'après la première question, on en déduit que $T \sim \text{Loi}(X|X > a)$. Ceci permet de simuler la loi de X conditionnellement à $X > a$ très simplement :

$$U \sim \text{Unif}([0, 1]) \implies T = F^{-1}(F(a) + (1 - F(a))U) \sim \text{Loi}(X|X > a).$$

Cette méthode est bien plus efficace que la méthode de rejet précédente, puisqu'on ne rejette rien et qu'elle fonctionne quel que soit a . Elle nécessite cependant la connaissance et l'inversion de F , contrairement à la méthode de rejet, qui n'exigeait que la capacité à simuler suivant la loi de X .

Remarques :

- Si F n'est pas bijective, tout ce qui précède est encore vrai en notant F^{-1} l'inverse généralisée de F (ou fonction quantile). Il suffit pour s'en convaincre d'utiliser la relation fondamentale vue en cours : $F(x) \geq u \Leftrightarrow x \geq F^{-1}(u)$.
- Bien entendu, cette méthode n'est rien d'autre que la méthode d'inversion...

4. Détailler comment simuler X de loi de Cauchy conditionnellement à $X > a$?

Correction Par l'exercice 1, la fonction de répartition d'une loi de Cauchy standard, de densité $x \mapsto 1/[\pi(1+x^2)]$, est

$$\forall x \in \mathbb{R} : F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x),$$

d'inverse généralisée :

$$\forall u \in [0, 1], F^{-1}(u) = \begin{cases} -\infty & \text{pour } u = 0 \\ \tan\left(\pi\left(u - \frac{1}{2}\right)\right) & \text{lorsque } u \in]0, 1[\\ \infty & \text{pour } u = 1. \end{cases}$$

En prenant $U \sim \text{Unif}([0, 1])$, on obtient comme générateur de $\text{Loi}(X|X > a)$:

$$F^{-1}(F(a) + (1 - F(a))U) = \tan\left(\arctan(a) + \left(\frac{\pi}{2} - \arctan(a)\right)U\right).$$

EXERCICE 7 ((*) Méthodes de Box-Muller et de Marsaglia)

On cherche ici à proposer deux méthodes de simulation de la loi $\mathcal{N}(0, 1)$.

1. Soit (U, V) un couple de variables i.i.d. de loi uniforme sur $[0, 1]$. Montrer que les variables X et Y définies par

$$\begin{cases} X = \sqrt{-2 \log U} \times \cos(2\pi V) \\ Y = \sqrt{-2 \log U} \times \sin(2\pi V) \end{cases}$$

sont i.i.d. gaussiennes centrées réduites. Le principe consistant à simuler U et V puis à calculer X et Y est appelé méthode de Box-Muller.

Indication (changement de variable multivarié) : Soient d un entier non-nul quelconque, U et V deux ouverts de $\mathbb{R}^d$, $f : V \rightarrow \mathbb{R}$ intégrable et $\varphi^{-1} : U \rightarrow V$ un $\mathcal{C}^1$ -difféomorphisme. Alors

$$\int_V f = \int_U f \circ \varphi^{-1} |\det(J_{\varphi^{-1}})|,$$

où $J_{\varphi^{-1}}$ est la matrice jacobienne de φ^{-1} .

Par exemple, dans le cas $d = 2$ où l'on cherche à évaluer

$$\int_V f(u, v) du dv,$$

il suffit d'écrire $(x, y) = \varphi(u, v)$, puis « $dx dy = |\det(J_{\varphi}(u, v))| du dv$ », ou $(u, v) = \varphi^{-1}(x, y)$ puis « $du dv = |\det(J_{\varphi^{-1}}(x, y))| dx dy$ », conduisant à

$$\int_V f(u, v) du dv = \int_U (f \circ \varphi^{-1})(x, y) \frac{dx dy}{|\det((J_{\varphi} \circ \varphi^{-1})(x, y))|} = \int_U (f \circ \varphi^{-1})(x, y) |\det(J_{\varphi^{-1}}(x, y))| dx dy.$$

On rappelle que, pour tout $(u, v) \in V$, si l'on écrit $\varphi(u, v) = (\varphi_1(u, v), \varphi_2(u, v)) \in \mathbb{R}^2$, alors :

$$J_\varphi(u, v) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial u}(u, v) & \frac{\partial \varphi_1}{\partial v}(u, v) \\ \frac{\partial \varphi_2}{\partial u}(u, v) & \frac{\partial \varphi_2}{\partial v}(u, v) \end{pmatrix},$$

et que φ est un $\mathcal{C}^1$ -difféomorphisme si et seulement si :

- (a) φ est une bijection ;
- (b) φ est $\mathcal{C}^1$;
- (c) pour tout $(u, v) \in V$, $\det(J_\varphi(u, v)) \neq 0$.

Correction On applique la méthode de la fonction muette. Soit donc $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction test (par exemple continue bornée), alors le Théorème de Transfert implique :

$$\begin{aligned} \mathbb{E}[h(X, Y)] &= \mathbb{E} \left[h \left(\sqrt{-2 \log U} \cos(2\pi V), \sqrt{-2 \log U} \sin(2\pi V) \right) \right] \\ &= \iint_{]0,1[^2} h \left(\sqrt{-2 \log u} \cos(2\pi v), \sqrt{-2 \log u} \sin(2\pi v) \right) \mathrm{d}v \mathrm{d}u. \end{aligned}$$

L'application $\varphi : (u, v) \mapsto (\sqrt{-2 \log u} \cos(2\pi v), \sqrt{-2 \log u} \sin(2\pi v))$ est clairement une bijection $\mathcal{C}^1$ de $]0,1[^2$ dans $\mathbb{R}^2$ (à des ensembles de mesure nulle près) et son jacobien vaut, pour tout $(u, v) \in]0,1[^2$:

$$\det(J_\varphi(u, v)) = \begin{vmatrix} \frac{-\cos(2\pi v)}{u\sqrt{-2 \log u}} & -2\pi\sqrt{-2 \log u} \sin(2\pi v) \\ \frac{-\sin(2\pi v)}{u\sqrt{-2 \log u}} & +2\pi\sqrt{-2 \log u} \cos(2\pi v) \end{vmatrix} = -\frac{2\pi}{u} \neq 0.$$

C'est donc bien un $\mathcal{C}^1$ -difféomorphisme et l'on peut appliquer le théorème de changement de variables (en notant $(x, y) = \varphi(u, v)$) :

$$\mathbb{E}[h(X, Y)] = \iint_{\mathbb{R}^2} h(x, y) \frac{1}{|\det(J_\varphi(\varphi^{-1}(x, y)))|} \mathrm{d}x \mathrm{d}y,$$

avec

$$\frac{1}{|\det(J_\varphi(\varphi^{-1}(x, y)))|} = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}},$$

puisque $\frac{1}{|\det(J_\varphi(u, v))|} = \frac{u}{2\pi}$ et $u = e^{-\frac{x^2+y^2}{2}}$. Cela donne au final :

$$\mathbb{E}[h(X, Y)] = \iint_{\mathbb{R}^2} h(x, y) \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} \mathrm{d}x \mathrm{d}y,$$

et montre que les variables X et Y sont bien i.i.d. gaussiennes centrées réduites.

2. Soit $(R, \theta) = (\sqrt{-2 \log U}, 2\pi V)$ les coordonnées polaires du point (X, Y) dans $\mathbb{R}^2$. Quelle est la loi du vecteur (R^2, θ) ?

Correction Par construction, R^2 et θ sont indépendantes et $\theta \sim \mathcal{U}([0, 2\pi])$. Le calcul de la fonction de répartition de R^2 montre ensuite que cette dernière suit une loi exponentielle de paramètre 1/2. Remarque : Pour la loi de R , une autre méthode (sans calcul) consiste à se souvenir que si U est uniforme sur $[0, 1]$ alors $-\log U \sim \text{Exp}(1)$. Ici, on a donc $R^2 = -2 \log U \sim \text{Exp}(1/2)$.

3. Déterminer la loi du couple de variables aléatoires $(W, Z) = (\sqrt{U} \cos(2\pi V), \sqrt{U} \sin(2\pi V))$. Comment peut-on simuler cette loi par méthode du rejet ?

Correction L'application $\varphi : (u, v) \in]0, 1]^2 \mapsto (\sqrt{u} \cos(2\pi v), \sqrt{u} \sin(2\pi v)) \in \mathcal{D}$, où $\mathcal{D}$ est le disque unité privé de son centre, est un C^1 -difféomorphisme, de jacobien π (informellement, le changement de variable $(w, z) = \varphi(u, v)$ aboutit donc à $dwdz = \pi \, \text{d}u \, \text{d}v$). Ainsi, pour toute fonction $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ continue bornée,

$$\mathbb{E}[h(W, Z)] = \iint_{]0, 1]^2} h(\sqrt{u} \cos(2\pi v), \sqrt{u} \sin(2\pi v)) \, \text{d}u \, \text{d}v = \iint_{\mathcal{D}} h(w, z) \frac{1}{\pi} \, \text{d}w \, \text{d}z.$$

Ceci montre que (W, Z) est uniformément distribué sur $\mathcal{D}$.

Remarque : Cela revient à déterminer la loi du couple de variables aléatoires

$$(W, Z) = \left(\frac{X}{\sqrt{X^2 + Y^2}} e^{-\frac{X^2 + Y^2}{4}}, \frac{Y}{\sqrt{X^2 + Y^2}} e^{-\frac{X^2 + Y^2}{4}} \right) = \left(\cos(\theta) e^{-\frac{R^2}{4}}, \sin(\theta) e^{-\frac{R^2}{4}} \right).$$

La densité de la loi que l'on cherche à échantillonner est $f : (w, z) \mapsto \frac{1}{\pi} \mathbf{1}_{(w, z) \in \mathcal{D}}$. Considérons la loi uniforme sur $[-1, 1]^2$, de densité $g : (w, z) \mapsto \frac{1}{4} \mathbf{1}_{(w, z) \in [-1, 1]^2}$. Nous avons bien $f \leq mg$, avec $m = \frac{4}{\pi}$, donc l'algorithme du rejet nous assure que $Z_\tau \sim \mathcal{U}(\mathcal{D})$ lorsque $(Z_n)_{n \geq 1}$ est i.i.d. selon $\mathcal{U}([0, 1]^2)$ et

$$\tau = \inf \{i \in \mathbb{N}^*, r(Z_i) \geq T_i \quad (\star)\},$$

où

- (a) $(T_n)_{n \geq 1}$ est i.i.d. selon $\mathcal{U}([0, 1])$ et indépendant de $(Z_n)_{n \geq 1}$;
- (b) $r(z) = \frac{f(z)}{mg(z)} = \frac{\mathbf{1}_{z \in \mathcal{D}}}{\mathbf{1}_{z \in [-1, 1]^2}}, \forall z \in \mathbb{R}^2$ (avec la possibilité de prendre la valeur ∞).

Dans notre cas, $r(Z_i) = \mathbf{1}_{Z_i \in \mathcal{D}}$ presque sûrement et le critère $(\star)$ est alors équivalent à $Z_i \in \mathcal{D}$ presque sûrement.

Autrement dit, l'algorithme du rejet revient à échantillonner sur le carré $[-1, 1]^2$ puis à conserver uniquement les points tombant à l'intérieur du disque unité.

4. Dédurre de la question précédente une nouvelle méthode de simulation de la loi $\mathcal{N}(0, 1)$. Cette dernière est appelée méthode polaire de Marsaglia. Quel est son avantage par rapport à l'algorithme de Box-Muller. Quel est son défaut ?

Correction Par construction,

$$\begin{cases} X &= \sqrt{-2 \log U} \cos(2\pi V) = \sqrt{-2 \log U} \frac{W}{\sqrt{U}} = \sqrt{-2 \log(W^2 + Z^2)} \frac{W}{\sqrt{W^2 + Z^2}} \\ Y &= \sqrt{-2 \log U} \sin(2\pi V) = \sqrt{-2 \log U} \frac{Z}{\sqrt{U}} = \sqrt{-2 \log(W^2 + Z^2)} \frac{Z}{\sqrt{W^2 + Z^2}}. \end{cases}$$

Pour échantillonner la loi $\mathcal{N}(0, 1)$, il suffit donc de simuler (W, Z) selon la méthode du rejet établie à la question précédente et d'appliquer la transformation.

L'avantage de la méthode polaire est de s'affranchir des calculs de cosinus et sinus (qui sont relativement coûteux) : en effet, le point $\left(\frac{W}{\sqrt{W^2 + Z^2}}, \frac{Z}{\sqrt{W^2 + Z^2}} \right)$ est sur le cercle unité, ainsi ses coordonnées peuvent naturellement être interprétées par $(\cos \theta, \sin \theta)$. Le défaut de cette approche est d'utiliser une méthode de rejet mais en réalité, on rejette relativement peu : en moyenne environ un point sur 5 (précisément une proportion égale à $1 - \pi/4 \approx 0,21$).

EXERCICE 8 (Loi de Poisson)

On s'intéresse à la simulation de la loi de Poisson : $\text{Poi}(\lambda)$, où $\lambda > 0$. Dans la bibliothèque Python `Numpy`, ceci est mis en place via deux méthodes décrites ci-dessous, selon la valeur de λ . La première génère concrètement un processus de Poisson, tandis que la seconde est une méthode d'inversion ne nécessitant pas de calculer l'inverse généralisée de la fonction de répartition de la loi $\text{Poi}(\lambda)$.

1. Méthode du processus de Poisson

(a) Soient (X_n) une suite de variables aléatoires i.i.d. de loi $\text{Exp}(\lambda)$ et

$$N = \sup \left\{ n \in \mathbb{N}, \sum_{k=1}^n X_k \leq 1 \right\}$$

(avec la convention qu'une somme vide vaut 0). Montrer que $N \sim \text{Poi}(\lambda)$.

Indication : Le volume du simplexe $\{x \in [0, 1]^n, x_1 + \dots + x_n \leq 1\}$ est $1/n!$

Correction N prend ses valeurs dans $\mathbb{N}$ et pour tout $n \in \mathbb{N}$ (en notant ci-dessous $\bar{x}_n = \sum_{k=1}^n x_k/n$),

$$\begin{aligned} \mathbb{P}(N = n) &= \mathbb{P} \left(\sum_{k=1}^n X_k \leq 1, \sum_{k=1}^{n+1} X_k > 1 \right) \\ &= \int_{\mathbb{R}_+^{n+1}} \mathbf{1}_{n\bar{x}_n \leq 1} \mathbf{1}_{n\bar{x}_n + x_{n+1} > 1} \lambda^{n+1} e^{-\lambda(n\bar{x}_n + x_{n+1})} dx_1 \dots dx_{n+1} \\ &= \int_{\mathbb{R}_+^n} \mathbf{1}_{n\bar{x}_n \leq 1} \lambda^n e^{-\lambda n\bar{x}_n} \left(\int_{\mathbb{R}_+} \mathbf{1}_{x_{n+1} > 1 - n\bar{x}_n} \lambda e^{-\lambda x_{n+1}} dx_{n+1} \right) dx_1 \dots dx_n \\ &= \int_{\mathbb{R}_+^n} \mathbf{1}_{n\bar{x}_n \leq 1} \lambda^n e^{-\lambda} dx_1 \dots dx_n \\ &= \frac{\lambda^n}{n!} e^{-\lambda}. \end{aligned}$$

Donc $N \sim \text{Poi}(\lambda)$.

Remarque : Si l'on connaît la théorie des processus de Poisson, ce résultat est classique. La variable aléatoire N compte le nombre d'événements jusqu'à la date $t = 1$, or de façon générale le nombre d'événements jusqu'à la date t suit une loi de Poisson de paramètre t .

(b) En déduire une manière de simuler la loi $\text{Poi}(\lambda)$ à partir de variables aléatoires uniformes.

Correction L'idée est de générer séquentiellement une suite de variables aléatoires i.i.d. de loi $\text{Exp}(\lambda)$ et de s'arrêter lorsque leur somme dépasse 1. Il suffit alors de compter combien on a pu « empiler de variables » avant de dépasser 1.

Ceci étant, d'après l'exercice 1, la loi $\text{Exp}(\lambda)$ peut être simulée à partir de variables aléatoires uniformes : soit (U_n) une suite de variables i.i.d. selon la loi $\text{Unif}([0, 1])$, alors $(-\log(U_n)/\lambda)$ est une suite de variables i.i.d. de loi $\text{Exp}(\lambda)$. Ensuite,

$$N \stackrel{\text{Loi}}{=} \sup \left\{ n \in \mathbb{N}, \sum_{k=1}^n -\log(U_k)/\lambda \leq 1 \right\} = \sup \left\{ n \in \mathbb{N}, \prod_{k=1}^n U_k \geq e^{-\lambda} \right\}.$$

Il suffit donc de calculer un produit cumulatif de variables aléatoires indépendantes et uniformes et de s'arrêter lorsque ce produit est strictement plus petit que $e^{-\lambda}$. La longueur du produit moins 1 suit la loi $\text{Poi}(\lambda)$.

2. Méthode du rejet transformée

- (a) Soient P une loi de fonction de répartition F et de densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}$, et $\phi :]0, 1[\rightarrow \mathbb{R}$ un C^1 -difféomorphisme. Montrer que $g = (f \circ \phi)|\phi'|$ (prolongée à 0 aux extrémités du segment $[0, 1]$) est une densité par rapport à la mesure de Lebesgue sur $[0, 1]$.

Correction Il est clair que $g \geq 0$ et le changement de variable $y = \phi(x)$ dans l'intégrale ci-dessous permet de conclure :

$$\int_{[0,1]} g(x) dx = \int_{[0,1]} f(\phi(x)) |\phi'(x)| dx = \int_{\mathbb{R}} f(y) dy = 1.$$

- (b) Soient Q la loi sur $[0, 1]$ de densité g et $Z \sim Q$. Quelle est la loi de $\phi(Z)$?

Correction Soit $h : \mathbb{R} \rightarrow \mathbb{R}$ une fonction continue bornée. Alors, le changement de variable $y = \phi(z)$ dans l'intégrale ci-dessous donne :

$$\mathbb{E}[h(\phi(Z))] = \int_{[0,1]} h(\phi(z)) f(\phi(z)) |\phi'(z)| dz = \int_{\mathbb{R}} h(y) f(y) dy.$$

Ainsi $\phi(Z) \sim P$.

- (c) On suppose qu'il existe $0 < m < \infty$ tel que $\|g\|_{\infty} \leq m$. Expliquer comment simuler la loi Q à partir de la loi uniforme.

Correction La loi Q étant sur $[0, 1]$, il suffit de prendre comme loi auxiliaire $\text{Unif}([0, 1])$. On a alors

$$\sup_{x \in [0,1]} \frac{g(x)}{1} \leq m,$$

et le ratio d'acceptation vaut, pour tout $x \in \mathbb{R}$, $g(x)/m$. Ainsi, on peut considérer deux suites indépendantes (Y_n) et (U_n) de variables aléatoires i.i.d. de loi $\text{Unif}([0, 1])$, et définir $N = \min\{n \geq 1, mU_n \leq f(\phi(Y_n))|\phi'(Y_n)|\}$. Il vient, par la méthode du rejet : $Y_N \sim Q$.

- (d) En pratique, ϕ a pour objectif d'approcher F^{-1} , de sorte qu'au lieu de simuler P par inversion classique, i.e. via $F^{-1}(U)$, où $U \sim \text{Unif}([0, 1])$, on la simule grâce à $\phi(Z)$, avec $Z \sim Q$ (on comprend que la loi Q est « proche » de la loi $\text{Unif}([0, 1])$).

Dans le but de simuler $\text{Poi}(\lambda)$, on choisit donc

$$f : x \in \mathbb{R} \mapsto \frac{\lambda^{\lfloor x \rfloor} \mathbf{1}_{x \geq 0}}{\lfloor x \rfloor!} e^{-\lambda} \mathbf{1}_{x \geq 0} \quad \text{et} \quad \phi : u \in]0, 1[\mapsto \frac{2}{1/2 - |u - 1/2|} (u - 1/2).$$

Montrer que f est bien une densité. On admettra que, dans ce cas, $\phi :]0, 1[\rightarrow \mathbb{R}$ est un C^1 -difféomorphisme et $\|g\|_{\infty} < \infty$.

Correction Il est clair que $f \geq 0$ et

$$\int_{\mathbb{R}} f(x) dx = \sum_{k \geq 0} \int_k^{k+1} f(x) dx = \sum_{k \geq 0} \frac{\lambda^k}{k!} e^{-\lambda} = 1.$$

Concernant la partie admise, il apparaît que ϕ ainsi choisie est dérivable sur $]0, 1[$ avec

$$\forall u \in]0, 1[, \quad \phi'(u) = \frac{1}{(1/2 - |u - 1/2|)^2},$$

donc, lorsque $u \neq 1/2$, on a

$$\phi'(u) = \frac{\phi(u)^2}{4(u - 1/2)^2}.$$

Ensuite,

$$\forall x \in]0, 1[, \quad g(x) = \frac{\lambda^{\lfloor \phi(x) \mathbf{1}_{\phi(x) \geq 0} \rfloor}}{\lfloor \phi(x) \mathbf{1}_{\phi(x) \geq 0} \rfloor!} e^{-\lambda} \mathbf{1}_{\phi(x) \geq 0} |\phi'(x)|,$$

et lorsque $x < 1/2$, $\phi(x) < 0$ donc $g(x) = 0$. De plus, lorsque $x \geq 3/4$, $\phi(x) \geq 0$ et

$$\begin{aligned} g(x) &= \frac{\lambda^{\lfloor \phi(x) \rfloor}}{\lfloor \phi(x) \rfloor!} e^{-\lambda} \frac{\phi(x)^2}{4(x - 1/2)^2} \\ &\leq 4 \frac{\lambda^{\lfloor \phi(x) \rfloor}}{\lfloor \phi(x) \rfloor!} e^{-\lambda} \lfloor \phi(x) + 1 \rfloor^2 \\ &\leq 4 \sum_{k \geq 0} \frac{\lambda^k}{k!} e^{-\lambda} (k + 1)^2 \\ &= \mathbb{E}[(X + 1)^2] < \infty, \end{aligned}$$

pour $X \sim \text{Poi}(\lambda)$.

- (e) En déduire un algorithme de simulation de la loi $\text{Poi}(\lambda)$.

Correction En reprenant la question 2c, on peut considérer deux suites indépendantes (Y_n) et (U_n) de variables aléatoires i.i.d. de loi $\text{Unif}([0, 1])$, et définir $N = \min\{n \geq 1, mU_n \leq f(\phi(Y_n))|\phi'(Y_n)|\}$ (avec m un majorant de ϕ). Par la méthode du rejet, $Y_N \sim Q$, donc $\phi(Y_N) \sim P$ et $\lfloor \phi(Y_N) \rfloor \sim \text{Poi}(\lambda)$. On remarque au passage que pour simuler suivant une loi discrète (ici $\text{Poi}(\lambda)$), on a en réalité seillé une simulation suivant la loi continue dont une densité est la prolongation constante par morceaux de la fonction de masse de la loi discrète considérée.

Remarque : il existe une façon de simuler la loi normale en se fondant sur la même idée. Dans ce cas, la fonction ϕ utilisée a la même forme que celle ci-dessus. Cela a du sens car, si l'on prend (X_n) i.i.d. selon $\text{Poi}(\lambda)$, alors par le TCL $\sum_{i=1}^n X_i \stackrel{\text{Loi}}{\approx} \mathcal{N}(n\lambda, n\lambda)$. Or par stabilité de la loi de Poisson par la somme, $\sum_{i=1}^n X_i \sim \text{Poi}(n\lambda)$. Autrement dit, pour $\bar{\lambda} > 0$ grand (en pratique plus grand que 5, Ce qui est le cas dans la situation d'utilisation de cette procédure de simulation – cf. la dernière question), $\text{Poi}(\bar{\lambda}) \approx \mathcal{N}(\bar{\lambda}, \bar{\lambda})$.

3. Justifier intuitivement qu'en pratique, la première méthode soit utilisée lorsque $\lambda < 10$ et la seconde lorsque $\lambda \geq 10$.

Correction La seconde méthode est très générale et fonctionne pour n'importe quelle valeur de λ . En revanche, elle fait appel à un mécanisme de rejet, ce qui la rend moins désirable que la première approche. Toutefois, dans la première méthode, lorsque λ est grand, les réalisations de la loi $\text{Exp}(\lambda)$ seront très proches de 0 (leur espérance est $1/\lambda$). Il faudra donc en « empiler » beaucoup pour dépasser 1, ce qui rendra la méthode peu efficace. 10 est bien sûr une valeur très arbitraire et non-optimisée.

EXERCICE 9 (★) Méthode du rejet sur l'hypographe d'une densité

Soit P une loi de densité f par rapport à la mesure de Lebesgue. On définit, pour toute fonction $h : \mathbb{R} \rightarrow \mathbb{R}_+$, son hypographe (au-dessus de l'axe des abscisses) par :

$$\text{hyp}(h) = \{(x, u) \in \mathbb{R} \times \mathbb{R}_+, u \leq h(x)\}.$$

1. Montrer que la surface de $\text{hyp}(f)$ est finie.

Correction Par intégration d'une fonction positive, on obtient :

$$\iint_{\mathbb{R}^2} \mathbf{1}_{(x,u) \in \text{hyp}(f)} dx du = \int_{\mathbb{R}} \left(\int_0^{f(x)} du \right) dx = \int_{\mathbb{R}} f(x) dx = 1.$$

2. Soit $(X, U) \sim \text{Unif}(\text{hyp}(f))$. Quelle est la loi marginale de X ?

Correction Le couple (X, U) a pour densité $(x, u) \mapsto \mathbf{1}_{(x,u) \in \text{hyp}(f)}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$. Il vient que X possède une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$, s'écrivant :

$$\forall x \in \mathbb{R}, \quad \int_{\mathbb{R}} \mathbf{1}_{(x,u) \in \text{hyp}(f)} du = f(x).$$

Ainsi $X \sim P$.

3. Soit maintenant une loi Q de densité g par rapport à la mesure de Lebesgue, telle qu'il existe $m > 0$ vérifiant : $\forall x \in \mathbb{R}, f(x) \leq mg(x)$.

Montrer que le couple (Y, U) dont la loi est définie par

$$Y \sim Q \quad \text{et} \quad U|Y \sim \text{Unif}([0, mg(Y)])$$

suit la loi uniforme sur $\text{hyp}(mg) : (Y, U) \sim \text{Unif}(\text{hyp}(mg))$.

Correction Soit $S = \{(y, u) \in \mathbb{R} \times \mathbb{R}_+, g(y) > 0\}$. Le support de $\text{Loi}(Y, U)$ est inclus dans S puisque $\mathbb{P}((Y, U) \notin S) = \mathbb{P}(g(Y) = 0) = 0$. Par ailleurs, par construction, le couple (Y, U) possède une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$, s'écrivant :

$$\forall (y, u) \in S : g(y) \frac{\mathbf{1}_{0 \leq u \leq mg(y)}}{mg(y)} = \frac{\mathbf{1}_{0 \leq u \leq mg(y)}}{m} = \frac{\mathbf{1}_{(y,u) \in \text{hyp}(mg)}}{\iint_{\mathbb{R}^2} \mathbf{1}_{(y',u') \in \text{hyp}(mg)} dy' du'}.$$

Ainsi $\text{Loi}(Y, U)$ a une densité qui coïncide sur son support avec la densité de la loi $\text{Unif}(\text{hyp}(mg))$. D'où $(Y, U) \sim \text{Unif}(\text{hyp}(mg))$.

4. En remarquant que $\text{hyp}(f) \subset \text{hyp}(mg)$, décrire une méthode de simulation de la loi P à partir de la loi Q (que l'on suppose savoir simuler facilement).

Correction On sait que si l'on simule $(X, U) \sim \text{Unif}(\text{hyp}(f))$, alors $X \sim P$. Ensuite, puisque $\text{hyp}(f) \subset \text{hyp}(mg)$, il est possible de simuler la loi $\text{Unif}(\text{hyp}(f))$ par rejet depuis $\text{Unif}(\text{hyp}(mg))$. Enfin, cette dernière loi peut être simulée selon le schéma décrit à la question 3.

Pour récapituler :

- (a) Simuler $Y \sim Q$ puis $U|Y \sim \text{Unif}([0, mg(Y)])$ (on obtient $(Y, U) \sim \text{Unif}(\text{hyp}(mg))$).
- (b) Si $U \leq f(Y)$, définir $X = Y$, sinon retourner à l'étape précédente.
- (c) Retourner X .

5. Si la densité f a pour support un intervalle borné et est telle que $\|f\|_\infty < \infty$, quelle loi Q proposeriez-vous ? Décrire l'algorithme résultant.

Correction Si f est bornée et a son support dans un intervalle borné, alors $\text{hyp}(f) \subset [a, b] \times [0, c]$, pour des réels a, b et c . En gardant à l'esprit la deuxième étape de la méthode proposée, il est alors possible de simuler $(Y, U) \sim \text{Unif}([a, b] \times [0, c])$. Dans ce cas, on peut choisir $Q = \text{Unif}([a, b])$ et l'algorithme est similaire à celui proposé à la question précédente en remplaçant la première étape par :

- (a) Simuler indépendamment $Y \sim \text{Unif}([a, b])$ et $U \sim \text{Unif}([0, c])$.

EXERCICE 10 ((†) Simulation de la loi Gamma)

Soient $a > 0$ et $\lambda > 0$. On souhaite simuler suivant la loi $\text{Gamma}(a, \lambda)$, dont une densité par rapport à la mesure de Lebesgue est donnée par :

$$f : x \mapsto \frac{\lambda^a x^{a-1}}{\Gamma(a)} e^{-\lambda x} \mathbf{1}_{x>0}.$$

On suppose dans la suite que $a > 1$ et on note

$$g : x \in \mathbb{R} \mapsto x^{a-1} e^{-x} \mathbf{1}_{x>0} \quad \text{et} \quad \mathcal{D} = \left\{ (x, y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*, x \leq \sqrt{g(y/x)} \right\}.$$

1. Calculer $\sup_{x \in \mathbb{R}} g(x)$ et $\sup_{x \in \mathbb{R}} x^2 g(x)$. En déduire que $\mathcal{D} \subset [0, x_a] \times [0, y_a]$, où $x_a = \left(\frac{a-1}{e}\right)^{\frac{a-1}{2}}$ et $y_a = \left(\frac{a+1}{e}\right)^{\frac{a+1}{2}}$.

Correction Soient $p \in \mathbb{N}$ et $h : x \in \mathbb{R} \mapsto x^p g(x)$. h est positive et nulle sur $\mathbb{R}_-$ donc $\sup_{x \in \mathbb{R}} h(x) = \sup_{x \in \mathbb{R}_+^*} h(x)$. Par ailleurs, h est dérivable sur $\mathbb{R}_+^*$, telle que

$$\forall x \in \mathbb{R}_+^*, \quad h'(x) > 0 \iff (a+p-1)x^{a+p-2}e^{-x} - x^{a+p-1}e^{-x} > 0 \iff x < a+p-1.$$

Or $a > 1$, donc $(a+p-1) \in \mathbb{R}_+^*$ et h atteint son maximum en $(a+p-1)$:

$$\sup_{x \in \mathbb{R}} h(x) = h(a+p-1) = \left(\frac{a+p-1}{e}\right)^{a+p-1}.$$

En particulier, pour $p = 0$ et $p = 2$, il vient $\sup_{x \in \mathbb{R}} g(x) = \left(\frac{a-1}{e}\right)^{a-1} = x_a^2$ et $\sup_{x \in \mathbb{R}} x^2 g(x) = \left(\frac{a+1}{e}\right)^{a+1} = y_a^2$.

Maintenant, pour tout $(x, y) \in \mathcal{D}$, on a trivialement

$$x \leq \sqrt{g(y/x)} \leq \sqrt{\sup_{x' \in \mathbb{R}} g(x')} = x_a$$

et

$$y^2 = \frac{y^2}{x^2} x^2 \leq \frac{y^2}{x^2} g\left(\frac{y}{x}\right) \leq \sup_{x' \in \mathbb{R}} x'^2 g(x') = y_a^2.$$

D'où $\mathcal{D} \subset [0, x_a] \times [0, y_a]$.

2. Soit $(X, Y) \sim \text{Unif}(\mathcal{D})$. Quelle est la loi de $W = Y/X$? En déduire que la surface de $\mathcal{D}$ vaut $\Gamma(a)/2$.

Correction On note $|\mathcal{D}|$ la surface de $\mathcal{D}$. Puisque $\mathcal{D} \subset [0, x_a] \times [0, y_a]$, $|\mathcal{D}| < \infty$ et il est licite de considérer la loi $\text{Unif}(\mathcal{D})$. Ceci étant, par construction, le couple (X, Y) a pour densité $(x, y) \in \mathbb{R}^2 \mapsto \mathbf{1}_{0 < x \leq \sqrt{g(y/x)}} \mathbf{1}_{y > 0} / |\mathcal{D}|$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$. Il vient que pour tout fonction $h : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée :

$$\mathbb{E}[h(W)] = \mathbb{E}\left[h\left(\frac{Y}{X}\right)\right] = \iint_{\mathbb{R}_+^* \times \mathbb{R}_+^*} h\left(\frac{y}{x}\right) \frac{1}{|\mathcal{D}|} \mathbf{1}_{x \leq \sqrt{g(y/x)}} dx dy.$$

Par le changement de variable $(u, v) = (x, y/x)$, i.e. $(x, y) = \varphi(u, v)$, avec $\varphi : (u, v) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \mapsto (u, uv) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$, qui est bien un C^1 -difféomorphisme, de jacobien $\det(J_\varphi(u, v)) = u \neq 0$ pour tout $(u, v) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$, il vient

$$\begin{aligned} \mathbb{E}[h(W)] &= \iint_{\mathbb{R}_+^* \times \mathbb{R}_+^*} h(v) \frac{1}{|\mathcal{D}|} \mathbf{1}_{0 < u \leq \sqrt{g(v)}} \mathbf{1}_{v > 0} |u| du dv \\ &= \int_{\mathbb{R}_+^*} h(v) \frac{1}{|\mathcal{D}|} \left(\int_0^{\sqrt{g(v)}} u du \right) dv \\ &= \int_{\mathbb{R}_+^*} h(v) \frac{1}{2|\mathcal{D}|} g(v) dv \\ &= \int_{\mathbb{R}} h(v) \frac{1}{2|\mathcal{D}|} v^{a-1} e^{-v} \mathbf{1}_{v > 0} dv. \end{aligned}$$

On reconnaît à côté de $h(v)$ le terme général de la loi $\text{Gamma}(a, 1)$. On en conclut que $W \sim \text{Gamma}(a, 1)$ et que $|\mathcal{D}| = \Gamma(a)/2$.

3. Conclure que $Z = W/\lambda \sim \text{Gamma}(a, \lambda)$.

Correction De nouveau par la méthode de la fonction muette, pour tout fonction $h : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée :

$$\mathbb{E}[Z] = \int_0^\infty h(w/\lambda) \frac{w^{a-1}}{\Gamma(a)} e^{-w} dw = \int_0^\infty h(z) \frac{\lambda^a z^{a-1}}{\Gamma(a)} e^{-\lambda z} dz,$$

par le changement de variable $z = w/\lambda$. Ainsi, $Z \sim \text{Gamma}(a, \lambda)$.

4. Comment simuler suivant les lois $\text{Unif}(\mathcal{D})$ et $\text{Gamma}(a, \lambda)$?

Correction Puisque $\mathcal{D} \subset [0, x_a] \times [0, y_a]$, on peut simuler suivant $\text{Unif}(\mathcal{D})$ via la méthode du rejet :

- (a) Simuler indépendamment $X \sim \text{Unif}([0, x_a])$ et $Y \sim \text{Unif}([0, y_a])$.
- (b) Si $X \leq \sqrt{g(Y/X)}$, retourner (X, Y) , sinon recommencer à l'étape précédente.

Concernant la loi $\text{Gamma}(a, \lambda)$, à partir de $(X, Y) \sim \text{Unif}(\mathcal{D})$, il suffit de calculer $Y/(\lambda X)$.

5. On vient de voir que pour simuler suivant la loi $\text{Gamma}(a, 1)$, il est inutile de connaître la constante $\Gamma(a)$ (seule la fonction g suffit). Est-ce que remplacer g par cg , où $c > 0$, dans la méthode ci-dessus change son efficacité ?

Correction L'efficacité de la méthode proposée ci-dessus réside dans le taux d'acceptation de l'algorithme du rejet pour simuler $\text{Unif}(\mathcal{D})$, qui vaut $1/m$ où m est le majorant choisi pour calculer le ratio d'acceptation. Ici (avec la convention $0/0 = 0$),

$$m = \sup_{(x,y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*} \frac{\mathbf{1}_{(x,y) \in \mathcal{D}} / |\mathcal{D}|}{\mathbf{1}_{(x,y) \in [0,x_a] \times [0,y_a]} / (x_a y_a)} = \frac{x_a y_a}{|\mathcal{D}|}.$$

Ainsi, la probabilité d'acceptation pour la méthode décrite ci-dessus vaut $\frac{|\mathcal{D}|}{x_a y_a} = \mathbb{P}((X,Y) \in \mathcal{D})$ lorsque $(X,Y) \sim \text{Unif}([0,x_a] \times [0,y_a])$. La question est donc de savoir si cette probabilité change en remplaçant g par cg .

Appelons $\mathcal{D}_c = \{(x,y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*, x \leq \sqrt{g(y/x)}\}$. En refaisant le raisonnement, il vient :

- (a) Pour tout $(x,y) \in \mathcal{D}_c$, $x \leq \sqrt{\sup_{x' \in \mathbb{R}} cg(x')} = \sqrt{cx_a}$ et $y^2 \leq \sup_{x' \in \mathbb{R}} x'^2 cg(x') = cy_a^2$. Donc $\mathcal{D}_c \subset [0, \sqrt{cx_a}] \times [0, \sqrt{cy_a}]$.
- (b) La surface de $\mathcal{D}_c$ vaut :

$$|\mathcal{D}_c| = \iint_{\mathbb{R}_+^* \times \mathbb{R}_+^*} \mathbf{1}_{0 < u \leq \sqrt{cg(v)}} \mathbf{1}_{v > 0} |u| du dv = \int_{\mathbb{R}_+^*} \frac{cg(v)}{2} dv = \frac{c\Gamma(a)}{2} \int_{\mathbb{R}} \frac{v^{a-1} e^{-v}}{\Gamma(a)} \mathbf{1}_{v > 0} dv = \frac{c\Gamma(a)}{2}.$$

Ainsi, simuler par rejet $\text{Unif}(\mathcal{D}_c)$ à partir de $\text{Unif}([0,x_a] \times [0,y_a])$ conduit à un taux d'acceptation de $\frac{|\mathcal{D}_c|}{(\sqrt{cx_a})(\sqrt{cy_a})} = \frac{|\mathcal{D}|}{x_a y_a}$. L'efficacité ne change donc aucunement avec c , et autant prendre $c = 1$.

Remarque : On aurait aussi pu remarquer que $(x,y) \in \mathcal{D}_c \iff (x/\sqrt{c}, y/\sqrt{c}) \in \mathcal{D}$, donc $\mathcal{D}_c$ est une dilatation (suivant les deux axes) de $\mathcal{D}$: $\mathcal{D}_c = \sqrt{c}\mathcal{D}$. Ainsi, il devient clair que $\mathcal{D}_c \subset \sqrt{c}([0,x_a] \times [0,y_a])$ et $|\mathcal{D}_c| = c|\mathcal{D}|$, et que la probabilité d'acceptation ne change pas.

EXERCICE 11 ((†) Algorithme Ziggourat)

Le but est de simuler selon une densité f de façon très efficace. Les deux idées-forces de la méthode Ziggourat sont élémentaires et ont déjà été croisées :

- Si le point (X,Y) est uniformément distribué dans l'hypographe de f , c'est-à-dire dans le domaine $\mathcal{H}(f) := \{(x,y) \in \mathbb{R} : 0 \leq y \leq f(x)\}$, alors X a pour densité f ;
- Si $f \leq g$ et si l'on sait simuler uniformément dans l'hypographe de g , alors on sait simuler uniformément dans $\mathcal{H}(f)$, donc selon f : il suffit d'effectuer des tirages uniformes dans $\mathcal{H}(g)$ jusqu'à tomber dans $\mathcal{H}(f)$, c'est la méthode de rejet basique.

Noter que g n'est pas une densité puisque $f \leq g$, mais peu importe. Le principe est ainsi de trouver une fonction g telle que :

- g majore f de façon très fine, afin de rejeter le moins possible ;
- il soit possible d'effectuer des tirages uniformes dans l'hypographe de g de façon très rapide.

Le cadre est le suivant : on considère une densité de probabilité $f : \mathbb{R} \rightarrow \mathbb{R}$ continue strictement décroissante sur $\mathbb{R}_+$ et nulle sur $\mathbb{R}_-^*$.

1. Première partie : Construction de g

On fixe un entier n , par exemple $n = 256$. Pour $x_1 > 0$, on note $y_1 = f(x_1)$ et

$$A = x_1 y_1 + \int_{x_1}^{\infty} f(y) dy,$$

puis par récurrence $y_{k+1} = y_k + A/x_k$ et $x_{k+1} = f^{-1}(y_{k+1})$. Le réel x_1 est choisi de telle sorte que $(x_n, y_n) = (0, f(0))$. Notons $\mathcal{T}_1 := \{(x,y) \in \mathbb{R}_+^2 : y \leq y_1 \mathbf{1}_{x < x_1} + f(x) \mathbf{1}_{x \geq x_1}\}$ la tranche inférieure et, pour $k = 2, \dots, n$, $\mathcal{T}_k := [0; x_{k-1}] \times [y_{k-1}; y_k]$ les tranches supérieures.

- (a) Faire un dessin pour comprendre comment est construite la séquence $((x_k, y_k))_{1 \leq k \leq n}$.

Correction La séquence $((x_k, y_k))_{1 \leq k \leq n}$ est construite de telle sorte que les tranches $\mathcal{T}_1, \dots, \mathcal{T}_n$ aient toutes la même surface, en l'occurrence égale à A .

- (b) Expliquer pourquoi une méthode de dichotomie permet de déterminer la séquence $((x_k, y_k))_{1 \leq k \leq n}$ avec une précision arbitraire.

Correction La fonction $A : \mathbb{R}_+ \rightarrow]0, 1]$ qui à $x > 0$ associe

$$A(x) := xf(x) + \int_x^\infty f(y)dy$$

est continue (car f l'est) et strictement décroissante puisque¹ si $0 < x < x'$ (un dessin rend les choses limpides) :

$$A(x) = xf(x) + \int_x^{x'} f(y)dy + \int_{x'}^\infty f(y)dy > xf(x') + \int_{x'}^{x'} f(y)dy + \int_{x'}^\infty f(y)dy = A(x').$$

La fonction $x \mapsto A(x)$ étant positive décroissante, elle admet une limite $L \geq 0$ en l'infini. Cette limite ne peut être que 0 : en effet, si on avait $L > 0$, alors comme

$$\int_x^\infty f(y)dy \xrightarrow{x \rightarrow \infty} 0,$$

on aurait $f(x) \sim L/x$ en l'infini, ce qui est impossible car l'intégrale de f serait alors infinie, or f est une densité. On a par ailleurs $A(0) = 1$.

Si l'on fixe $x_1 > 0$, ceci fixe $y_1 = f(x_1)$ et $A = A(x_1)$. Ceci détermine $y_2 = y_1 + A/x_1$. Si $y_2 < f(0)$, alors il existe $x_2 \in]0, x_1[$ tel que $x_2 = f^{-1}(y_2)$, ce qui permet de déterminer $y_3 = y_2 + A/x_2$, et ainsi de suite. Puisque $x \mapsto A(x)$ effectue une bijection de $\mathbb{R}_+$ vers $]0, 1]$, on voit que si la condition initiale x_1 est trop proche de 0, il existera $k < n$ tel que $y_k \geq f(0)$, donc ça ne marchera pas. A l'inverse, si x_1 est trop grand, alors on aura $x_n > 0$, donc ça ne marche pas non plus. En résumé, la fonction $x_1 \mapsto x_n$ est ou bien non définie (s'il existe $k < n$ tel que $y_k \geq f(0)$) ou bien strictement croissante. Par conséquent il n'existe qu'une valeur possible x_1 telle que $(x_n, y_n) = (0, f(0))$ et celle-ci peut être déterminée par dichotomie.

Exemple : pour la demi-gaussienne, i.e. $f(x) \propto e^{-\frac{x^2}{2}} \mathbf{1}_{x \geq 0}$, et $n = 256$ tranches on trouve

$$x_1 \approx 3.6541528853610088 \quad \text{et} \quad A \approx 0.00492867323399$$

- (c) Cette étape préliminaire est bien sûr coûteuse mais elle est effectuée **une seule fois** et permet d'aboutir à la fonction majorante g en escalier (hormis pour la rampe d'accès à droite de x_1) :

$$g(x) := f(x) \mathbf{1}_{x \geq x_1} + \sum_{k=2}^n y_k \mathbf{1}_{x_k \leq x < x_{k-1}}.$$

Si l'on suppose savoir tirer uniformément dans $\mathcal{T}_1$ (nous reviendrons sur ce point ultérieurement), expliquer comment simuler uniformément dans l'hypographe de g .

Correction Puisque toutes les tranches ont la même surface A , il suffit de choisir une tranche au hasard puis d'effectuer un tirage uniforme dans cette tranche :

1. Pour f dérivable c'est trivial car alors la fonction A l'est aussi avec $A'(x) = xf'(x) < 0$ car f est strictement décroissante.

- i. simuler $K \sim \text{Unif}(\{1, \dots, n\})$;
- ii. si $K = 1$, on tire uniformément le point (X, Y) dans $\mathcal{T}_1$ par une méthode ad hoc ; si $2 \leq K \leq n$, il suffit de tirer $X \sim \text{Unif}([0; x_{k-1}])$ et $Y \sim \text{Unif}([y_{k-1}; y_k])$. Nous verrons plus loin qu'on n'a pas forcément besoin de simuler Y , mais passons.

2. Deuxième partie : Méthode de rejet

Nous savons maintenant simuler (X, Y) uniformément dans l'hypographe de g à partir du tirage uniforme dans l'une des tranches. Comme nous allons le voir, une astuce permet très souvent de se dispenser du tirage de Y . Pour l'instant, on suppose donc uniquement $K \sim \text{Unif}(\{1, \dots, n\})$.

- (a) Vérifier que si $2 \leq K = k \leq n$ et $X \sim \text{Unif}([0; x_{k-1}])$, alors si $X \leq x_k$ on accepte X .

Correction A priori on devrait simuler $Y \sim \text{Unif}([y_{k-1}; y_k])$ et vérifier si $Y \leq f(X)$ mais c'est inutile car si $X \leq x_k$, alors $f(X) \geq f(x_k) = y_k \geq Y$. Bilan : nul besoin de simuler Y , la proposition X est automatiquement acceptée!

- (b) On suppose toujours $2 \leq K \leq n$ et $X \sim \text{Unif}([0; x_{k-1}])$, mais cette fois $x_k < X \leq x_{k-1}$. Comment appliquer l'idée du rejet ?

Correction Cette fois il faut simuler $Y \sim \text{Unif}([y_{k-1}; y_k])$ puis : si $Y \leq f(X)$, on accepte X ; si $Y > f(X)$, on refuse X et on revient au début, c'est-à-dire au tirage de $K \sim \text{Unif}(\{1, \dots, n\})$.

- (c) On suppose désormais $K = 1$ et on note $x_0 := A/y_1$. Montrer que si $X \sim \text{Unif}([0; x_0])$, alors si $X < x_1$, on accepte X .

Correction Il faut tirer uniformément dans $\mathcal{T}_1$. Ecrivons $\mathcal{T}_1 = \mathcal{T}'_1 \cup \mathcal{T}''_1$ avec $\mathcal{T}'_1 := [0; x_1[\times [0; y_1[$ et $\mathcal{T}''_1 = \mathcal{T}_1 \setminus \mathcal{T}'_1$. Si on note μ_1, μ'_1 et μ''_1 les lois uniformes sur ces trois ensembles, on peut voir μ_1 comme un mélange des lois μ'_1 et μ''_1 :

$$\mu_1 = \frac{|\mathcal{T}'_1|}{|\mathcal{T}_1|} \mu'_1 + \frac{|\mathcal{T}''_1|}{|\mathcal{T}_1|} \mu''_1,$$

ce qui signifie qu'on peut procéder en deux temps : on fait un premier tirage pour décider du domaine, puis on simule uniformément dans ce domaine. Pour le premier tirage, il suffit de connaître les pondérations, or

$$\frac{|\mathcal{T}'_1|}{|\mathcal{T}_1|} = \frac{x_1 y_1}{A} = \frac{x_1 y_1}{x_0 y_1} = \frac{x_1}{x_0}.$$

Or $X \sim \text{Unif}([0; x_0])$ vérifie $X \leq x_1$ avec probabilité x_1/x_0 donc dans ce cas on devrait simuler un point (X', Y') uniformément dans $\mathcal{T}'_1 := [0; x_1[\times [0; y_1[$, mais c'est inutile : on dispose déjà de X uniforme sur $[0; x_1[$ et, pour les mêmes raisons que ci-dessus, il n'y a pas besoin de simuler Y' puisqu'on aura forcément $Y' \leq f(X)$. Bref, on accepte X .

- (d) Si le tirage précédent a donné $X \geq x_1$, montrer qu'il "suffit" de générer un nouveau X selon la densité $cf(x)\mathbf{1}_{x \geq x_1}$.

Correction Vu ce qui précède, la probabilité que $X \sim \text{Unif}([0; x_0])$ vérifie $X \geq x_1$ est exactement celle avec laquelle on doit simuler un point uniformément dans

$$\mathcal{T}''_1 = \{(x, y) \in \mathbb{R}_+^2 : x \geq x_1, y \leq f(x)\}.$$

Cependant, n'oublions pas que si l'on y arrive, i.e. si $(X, Y) \sim \text{Unif}(\mathcal{T}_1'')$, alors on conservera uniquement X . Par conséquent, il suffit de montrer que dans ce cas X a pour densité $cf(x)\mathbf{1}_{x > x_1}$. Or cette vérification est directe car

$$|\mathcal{T}_1''| = \int_{x_1}^{\infty} f(x)dx = \frac{1}{c},$$

et pour toute fonction test φ , si $(X, Y) \sim \text{Unif}(\mathcal{T}_1'')$, il découle de Fubini que

$$\mathbb{E}[\varphi(X)] = \iint_{\mathbb{R}^2} \varphi(x) \frac{\mathbf{1}_{\mathcal{T}_1''}(x, y)}{|\mathcal{T}_1''|} dx dy = \int_{\mathbb{R}} \varphi(x) cf(x) \mathbf{1}_{x > x_1} dx.$$

-
- (e) Il reste maintenant à expliquer comment générer X selon la densité $cf(x)\mathbf{1}_{x \geq x_1}$. Ceci dépend bien sûr de f . Si f est la loi exponentielle de paramètre 1, montrer que c'est élémentaire. Si f est la densité de la variable absolue de la gaussienne standard, montrer qu'on peut se servir de la loi $x_1 + \text{Exp}(x_1)$ comme loi de proposition et appliquer une méthode de rejet.

Correction Par la propriété d'absence de mémoire, la loi d'une variable exponentielle sachant qu'elle est supérieure à $x_1 > 0$ est la loi exponentielle translatée de x_1 donc la simulation est immédiate.

Si f est la densité de la variable absolue de la gaussienne standard, alors la densité d'une variable ayant cette loi sachant qu'elle est supérieure à $x_1 > 0$ est

$$f_1(x) = c_1 e^{-\frac{x^2}{2}} \mathbf{1}_{x \geq x_1},$$

tandis que celle de la loi instrumentale $x_1 + \text{Exp}(x_1)$ est

$$g(x) = x_1 e^{-x_1(x-x_1)} \mathbf{1}_{x \geq x_1}.$$

Dans ce cas, un calcul rapide montre que

$$m := \sup_{x \geq x_1} \frac{f_1(x)}{g(x)} = \frac{f_1(x_1)}{g(x_1)},$$

donc, pour tout $x \geq x_1$, le rapport d'acceptation vaut

$$r(x) := \frac{f_1(x)}{mg(x)} = e^{-\frac{1}{2}(x-x_1)^2}.$$

Au total, pour simuler selon f_1 , on simule des variables $X \sim x_1 + \text{Exp}(x_1)$ et des variables uniformes sur $[0, 1]$ jusqu'à ce que

$$U \leq r(X) = e^{-\frac{1}{2}(X-x_1)^2}.$$

Dès que cette condition est satisfaite, l'algorithme renvoie X .

3. Troisième partie : Synthèse

On se focalise sur la loi normale standard en supposant que la séquence $0 = x_n < \dots < x_0$ est donnée une fois pour toutes.

- (a) Pourquoi n'est-il pas restrictif d'avoir supposé $f = 0$ sur $\mathbb{R}_-^*$?

Correction La densité de $X \sim |\mathcal{N}(0, 1)|$ est

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathbf{1}_{x \geq 0}$$

et répond aux conditions d'application précédentes puisqu'elle est continue strictement décroissante sur $\mathbb{R}_+$. De plus si $X \sim f$ et si R suit une loi de Rademacher et est indépendante de X , alors $W := RX \sim \mathcal{N}(0, 1)$.

-
- (b) Donner le pseudo-code de l'algorithme Ziggourat.

Correction L'algorithme peut donc se décrire comme suit :

- 1 Simuler $K \sim \text{Unif}(\{1, \dots, n\})$.
- 2 Simuler $U \sim \text{Unif}([0, 1])$, calculer $X = x_{K-1}U$ et :
 - Si $X < x_K$, renvoyer X .
 - Sinon :
 - Si $2 \leq K \leq n$, simuler $V \sim \text{Unif}([0, 1])$, calculer $Y = y_{K-1} + (y_K - y_{K-1})V$ et si $Y < f(X)$, renvoyer X , sinon retourner en 1.
 - Si $K = 1$, retourner X selon la densité $cf(x)\mathbf{1}_{x \geq x_1}$ selon une méthode adaptée (cf. Question 2e).

Remarque : on prendra garde au point suivant : si $K = 1$ et $x_0U \geq x_1$, on applique la méthode de rejet jusqu'à obtenir $X \sim cf(x)\mathbf{1}_{x \geq x_1}$, autant de fois que nécessaire ; tandis que si $K \geq 2$ et $x_{K-1}U \geq x_K$, on ne teste $Y < f(X)$ qu'une seule fois, c'est-à-dire que si $Y \geq f(X)$ il faut retourner au début et tirer $K \sim \text{Unif}(\{1, \dots, n\})$, et non rester dans la même tranche jusqu'à ce que la condition $Y < f(X)$ soit satisfaite.

-
- (c) Discuter de l'efficacité de cet algorithme.

Correction La seule étape coûteuse est le tirage selon la densité $cf(x)\mathbf{1}_{x \geq x_1}$, mais on n'y fait appel que moins d'une fois sur n , puisque même lorsque le premier tirage donne $K = 1$, alors si $X = x_0U < x_1$, on n'y a pas recours. De plus, quelle que soit la tranche, la plupart du temps on ne fait pas appel à f car la condition $x_{K-1}U < x_K$ est très souvent satisfaite lorsque n est grand (les x_k étant alors proches les uns des autres).

-
- (d) Pourquoi l'algorithme Ziggourat porte-t-il ce nom ?

Correction Ce nom vient de la forme de la fonction g , qui rappelle celle des [édifices religieux mésopotamiens](#). Cet algorithme est dû à George Marsaglia et Wai Wan Tsang, *The Ziggurat Method for Generating Random Variables*, Journal of Statistical Software, Vol. 5(8), 2000.
