# Cours/TP - Agrégation externe de mathématiques - Simulations

Référence : Ycart, P.S. Toulouse

# 1 Simulations

## 1.1 Simulation de lois de Poisson

Une v.a. discrète $X$ suit une loi de Poisson de paramètre $\alpha > 0$ si

$$P(X = k) = \frac{e^{-\alpha}\alpha^k}{k!}$$

Nous allons proposer un algorithme de simulation de $X$ différent de celui d'inversion de la fonction de répartition.

**Exercice 1.1** 1. *Soit $E_i$ une suite de v.a. iid de loi exponentielle $\alpha$. On note*

$$S_n = E_1 + \cdots + E_n$$

*Montrer qu'on a :*

$$P(S_n \leq 1 < S_{n+1}) = \frac{e^{-\alpha}\alpha^n}{n!}$$

2. *En conclure un algorithme de simulation de lois de Poisson de paramètre $\alpha$.*
3. *Quel est le nombre moyen de variables uniformes nécessaires pour simuler cette loi de Poisson ?*
4. *Donner à 99% un encadrement du nombre de lois uniformes nécessaires pour générer 400 lois de Poisson de paramètre $\alpha = 4$.*

## 1.2 Simulation par méthode de rejet de lois à densité connue

### Méthode générale

La méthode du rejet permet de simuler des v.a. ou des vecteurs aléatoires dont la loi est à densité par rapport à la mesure de Lebesgue, en s'aidant d'une autre densité dont on sait simuler la loi. Voici l'algorithme de simulation. On souhaite simuler Z dans $\mathbb{R}^d$ de densité $f$. On suppose que l'on sait simuler $X$ de loi $g$ et trouver $0 < c < 1$ tel que $f \leq cg$.
- Générer $X_i \sim g$ et $U_i \sim \mathcal{U}_{[0;1]}$.
- Calculer $M_i = (X_i, cg(X_i)U_i)$
- Dès que $cg(X_{i_0})U_{i_0} \leq f(X_{i_0})$, renvoyer $Z = X_{i_0}$, sinon faire $i + 1 -> i$.

Le vecteur aléatoire $Z$ a alors pour densité $f$.

**Exercice 1.2** 1. *Soit $f$ densité sur $\mathbb{R}^d$ et $G$ l'hypographe :*

$$G = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} \qquad 0 \leq y \leq f(x)\}$$

*On pose $M(Z, Y)$ de loi uniforme sur $G$. Montrer qu'alors la loi du vecteur aléatoire $Z$ a pour densité $f$ par rapport à la mesure de Lebesgue $\lambda_d$.*

2. *Soit $X$ vecteur aléatoire de densité $g$ par rapport à $\lambda_d$ et $M = (X, cg(X)U)$ où $U$ est une v.a. uniforme sur $[0; 1]$, indépendante de $X$ et $c > 0$. On note $H$ l'hypographe :*

$$H = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} \qquad 0 \leq y \leq cg(x)\}$$

*Montrer que $M$ suit une loi uniforme sur $H$.*

3. *Conclure.*

**Application à la loi Gamma et loi de Weibull.**

Écrire un code Matlab simulant une loi Gamma de paramètre $a > 0$ quelconque donnée par la densité :

$$f_a(t) = \frac{1}{\Gamma(a)} t^{a-1} e^{-t} \chi_{t>0}$$

en utilisant l'algorithme précédent et la loi de Weibull de paramètre $a$ de densité $g$ donnée par :

$$g(t) = at^{a-1} e^{-t^a} \chi_{t>0}$$

## 1.3 Non vieillissement de la loi exponentielle

1. Montrer que si $X$ est une variable aléatoire de loi exponentielle, alors on a :

$$P(X \leq x + s | X \geq s) = P(X \leq x)$$

2. Illustrer le résultat précédent : faire $N$ simulations d'une loi exponentielle et tracer sur un même graphique la fonction densité de la loi exponentielle et l'histogramme empirique de « $X - s$ sachant $X \geq s$ ».

## 1.4 Échantillonnage préférentiel

### Méthode générale

On cherche à estimer, à partir d'une variable aléatoire $X$ de densité connue, des quantités du type $\mathbb{E}[f(X)]$. L'idée générale est d'utiliser la convergence presque sure et dans $L^1$ de la moyenne empirique d'un échantillon de v.a. i.i.d. vers leur espérance commune. Cependant, en terme de variance de l'estimation donnée par cette moyenne empirique, il peut être intéressant d'utiliser un $n$-échantillon de $\tilde{X}$, autre v.a. à densité par rapport à $X$.

**Exercice 1.3** *Démontrer que si $\tilde{X}$ possède une densité par rapport à $X$, et que $(\tilde{X}_1, \ldots, \tilde{X}_n)$ désigne un $n$-échantillon, alors*

$$\frac{1}{n} \sum_{i=1}^{n} f(\tilde{X}_i) \frac{P_X(\tilde{X}_i)}{P_{\tilde{X}}(\tilde{X}_i)} \longrightarrow \mathbb{E}[f(X)]$$

### Application à l'estimation des moments de $\mathcal{N}(0, 1)$

**Exercice 1.4**

1. *Utiliser la propriété précédente pour $X \sim \mathcal{N}(0, 1)$ et $\tilde{X} \sim \mathcal{L}(\lambda)$. Calculer*

$$\tilde{Y} = f(\tilde{X}) \frac{P_X(\tilde{X})}{P_{\tilde{X}}(\tilde{X})}$$

*ainsi que la variance des estimateurs obtenus par le biais de $\tilde{X}$.*

2. *Comparer avec la variance de l'estimateur donné par $X$ lorsque $f(u) = u$ pour $\lambda = 1$. (Code Matlab)*
3. *Définir le "meilleur" $\lambda$ et en donner une estimation via Matlab.*

### Application à l'estimation de la queue d'une gaussienne $\mathcal{N}(0, 1)$

On souhaite désormais estimer la probabilité

$$P\left(\mathcal{N}(0, 1) \geq 3\right)$$

**Exercice 1.5** *On définit la variable aléatoire $Y$ par $Y = \chi_{X \geq 3}$.*

1. *Quelle est la loi de* $Y$ *?*
2. *On définit* $\tilde{X} \sim \mathcal{N}(3,1)$ *et* $\tilde{Y} = \chi_{\tilde{X} \geq 3}$.
3. *Quel est l'estimateur par la méthode de l'échantillonage préférentiel ici ?*
4. *Calculer la variance des deux estimateurs (ceux donnés par* $X$ *et* $\tilde{X}$*).*
5. *Donner une approximation (en utilisant Matlab) de la "meilleure" loi* $\mathcal{N}(m,1)$ *pour* $\tilde{X}$.

## 1.5 Loi des grands nombres

**Exercice 1.6**

1. *Faire* $N$ *simulations* $X_1, \ldots X_N$ *d'une loi exponentielle de paramètre* $\lambda$ *et tracer sur un même graphique la fonction qui à* $N$ *associe la moyenne empirique. Qu'observe-t-on ?*
2. *Faire de même pour une loi de Cauchy et vérifier expérimentalement que la loi des grands nombres ne s'applique pas dans ce cas-là.*

## 1.6 Mouvement Brownien

Créer un code Matlab simulant une trajectoire brownienne.