M1 MAPI3 - Simulations stochastiques 2024-2025
Xavier Bressaud *xavier.bressaud@math.univ-toulouse.fr*
Manon Costa *manon.costa@math.univ-toulouse.fr*
Bastien Mallein *bastien.mallein@math.univ-toulouse.fr*

TD 1 & TP 1 - Simulation d'une variable aléatoire

Lois discrètes à support fini.
Soit $x_1, x_2, \cdots, x_n$ des nombres réels tous différents et soit $p_1, p_2, \cdots, p_n$ des nombres réels positifs tels que $\sum_{i=1}^n p_i = 1$. On pose $s_0 = 0$ et pour tout $1 \leq k \leq n$, $s_k = \sum_{i=1}^k p_i$. Soit U est une variable aléatoire de loi uniforme $\mathcal{U}([0,1])$ et

$$X = \sum_{k=1}^n x_k \mathbf{1}_{(s_{k-1} \leq U \leq s_k)}.$$

Alors, X est une variable aléatoire de loi discrète $P = p_1\delta_{x_1} + p_2\delta_{x_2} + \cdots + p_n\delta_{x_n}$.

Exercice 1.

1. Comment simuler une variable aléatoire X $\{1,2,3\}$, de loi donnée par $\mathbb{P}(X=1) = 0,3$, $\mathbb{P}(X=2) = 0,1$ et $\mathbb{P}(X=3) = 0,6$ à partir d'un tirage uniforme sur $[0,1]$.
2. **TP -** Ecrire une fonction générique permettant de simuler une variable aléatoire discrète à partir du vecteur $(x_1, \ldots x_d)$ des valeurs possibles et du vecteur $(p_1, \ldots, p_d)$ des probabilités associées.
3. **TP -** L'appliquer au cas précédent et vérifier sur un histogramme les variables simulées ont la bonne loi.

Exercice 2.

1. Expliquer comment générer, à partir de variables $i.i.d$ uniformes sur $[0,1]$, une variable aléatoire X de loi Binomiale $\mathcal{B}(n,p)$ où les valeurs $n \geq 1$ et $0 < p < 1$ sont affectées par l'utilisateur.
2. **TP -** En utilisant la méthode précédente, coder une fonction binom(n,p,m) qui génère un vecteur de taille m de variables aléatoires binomiales indépendantes de paramètres (n,p).

Méthode par Inversion de la fonction de réparition. Soit X une variable aléatoire réelle de fonction de répartition F. On appelle inverse généralisée de F, la fonction G définie pour tout $y \in]0,1]$ par $G(y) = \inf\{x \in \mathbb{R} \ / F(x) \geq y\}$. Si U est une variable aléatoire de loi uniforme $\mathcal{U}([0,1])$, alors $G(U)$ a même loi que X.

Exercice 3.

1. Appliquer la méthode d'inversion de la fonction de répartition pour simuler une variable aléatoire `Y` de loi de Weibull de densité $3x^2e^{-x^3}$ sur $\mathbb{R}^+$.
2. **TP -** Écrire un code pour générer N réalisations indépendantes de la variable aléatoire `Y` de loi de Weibull de densité $3x^2e^{-x^3}$ sur $\mathbb{R}^+$.
3. Appliquer la méthode d'inversion de la fonction de répartition pour simuler une variable aléatoire `Z` de loi de Cauchy $\mathcal{C}(c)$ (avec $c > 0$), de densité $\frac{c}{\pi(c^2+x^2)}$. (On rappelle qu'arctan(x) est une primitive de $\frac{1}{(1+x^2)}$).
4. **TP -** Écrire un code pour générer N réalisations indépendantes de la variable aléatoire `Z` de loi de Cauchy $\mathcal{C}(c)$.
5. **TP -** Tracer les moyennes empiriques successives de `Y` et vérifier qu'elles convergent presque sûrement.
6. **TP -** Que se passe-t-il pour les moyennes empiriques successives de `Z` ?
7. **TP -** A l'aide d'un histogramme, observer quelle est la loi de la moyenne empirique associée à `Z` ?
8. Appliquer la méthode d'inversion de la fonction de répartition pour simuler une variable aléatoire `A` de loi de l'arcsinus, de densité $\frac{1}{\sqrt{\pi x(1-x)}}$ sur $[0,1]$.

Méthode par Troncature.

Exercice 4. Soit X une v.a. réélle positive à densité, de fonction de répartition F, et N est une variable aléatoire à valeurs dans $\mathbb{N}$ telle que, pour tout $n \in \mathbb{N}$, $\mathbb{P}[N = n] = F(n) - F(n-1)$.

1. Montrer que pour que N soit une variable à valeurs dans $\mathbb{N}$, necessairement X prend ses valeurs dans $[-1, +\infty[$.
2. Montrer que la partie entière $\lfloor X \rfloor + 1$ a même loi que N.
3. Vérifier qu'on peut l'utiliser pour simuler une variable aléatoire de loi uniforme $\mathcal{U}(\{1, 2, \cdots, n\})$ à partir d'une uniforme sur $[0, n]$; une loi géométrique de paramètre p à partir d'une exponentielle.

Calcul de loi.

Exercice 5. Calculer la loi de $\sqrt{-\ln(U)}$, si $U \sim \mathcal{U}nif(0,1)$.
En multipliant par une constante (à déterminer), en déduire une façon de simuler une loi de Rayleigh, de densité $\frac{x}{\sigma^2} e^{-x^2/2\sigma^2}$ sur $\mathbb{R}^+$.

Exercice 6. En faisant un mélange de loi, expliquer comment simuler une variable aléatoire de loi donnée par $\frac{1}{3}\mu_1 + \frac{2}{3}\mu_2$ avec μ_1 une loi exponentielle et μ_2 une uniforme sur $[0, 1]$.

Méthode du rejet.

Exercice 7. On souhaite simuler une variable aléatoire uniforme sur l'intérieur B d'une cardioïde. La cardioïde de paramètre $a > 0$ est la courbe fermée de $\mathbb{R}^2$ d'équation cartésienne

$$(x^2 + y^2 - ax)^2 = a^2(x^2 + y^2).$$

On peut également la définir par l'équation polaire $r(\theta) = a(1 + \cos(\theta))$.

1. Vérifier que le point $\Big(r(\theta)\cos(\theta), r(\theta)\sin(\theta)\Big)$ est située sur la cardioïde, et que le pavé $[-2a, 2a]^2$ contient la cardioïde.
2. En déduire un moyen de simuler un point uniformément sur l'intérieur de la cardioïde
$$B = \{(x, y) \in \mathbb{R}^2 : (x^2 + y^2 - ax)^2 \le a^2(x^2 + y^2)\}.$$
3. **TP -** Tracer la cardioïde, le pavé la contenant, et afficher un échantillon de taille n de loi uniforme sur B.
4. En utilisant une methode de Monte-Carlo, estimez l'aire contenue dans la cardioïde ainsi que la probabilité de rejet.
5. **TP -** Comment diminuer la probabilité de rejet ?

Exercice 8. Proposer un algorithme de rejet pour générer une variable aléatoire de densité $f(x) = \frac{\pi}{2}\sin(\pi x)$ sur $[0, 1]$. Quelle est la probabilité de rejet ?

Exercice 9. On veut simuler une variable aléatoire X de densité $f(x) = \frac{1}{Z}e^{-x^3}$ sur $\{x \in \mathbb{R} : x \ge 1\}$ à partir de la simulation d'une variable aléatoire Y de densité $g(x) = \frac{1}{x^2}$ sur $\{x \in \mathbb{R} : x \ge 1\}$. La constante de normalisation $Z = \int_1^{+\infty} e^{-x^3} dx$ n'est pas calculable, mais on n'en aura pas besoin pour les simulations.

1. Montrer qu'il est facile de simuler Y par l'inversion de la fonction de répartition.
2. Montrer que $x \mapsto \frac{f(x)}{g(x)}$ est décroissante sur $\{x \in \mathbb{R} : x \ge 1\}$.
3. En déduire que $f(x) \le \frac{1}{eZ} g(x)$ pour tout $x \ge 1$.
4. En déduire un algorithme de type rejet pour simuler X à partir de réalisations indépendantes de Y.

Gaussiennes : méthode de Box-Muller

Si U et V sont deux variables aléatoires indépendantes de loi uniforme $\mathcal{U}([0,1])$, alors $X = \sqrt{-2\,\log U}\cos(2\pi V)$ et $Y = \sqrt{-2\,\log U}\sin(2\pi V)$ sont indépendantes et de loi normale $\mathcal{N}(0,1)$.

Exercice 10.

1. Utiliser l'algorithme de Box-Muller pour générer N réalisations de variables aléatoires indépendantes et de loi normale $\mathcal{N}(m, s^2)$ où la moyenne $m \in \mathbb{R}$ et la variance $s^2 > 0$ sont affectées par l'utilisateur.
2. Proposer un algorithme de rejet pour générer une variable aléatoire standard en utilisant la méthode du rejet à partir de la loi de Cauchy de densité $\frac{1}{\pi(1+x^2)}$. Quelle est la probabilité de rejet ?
3. **TP -** Coder les deux méthodes proposées précédemment et vérifier par des histogrammes que les variables simulées suivent bien une loi gaussienne.
4. **TP -**Proposez une troisième méthode (plus approximative) utilisant le théorème central limite.
5. **TP -**En utilisant la fonction time() de Python, comparer les performances de ces trois manières de simuler des gaussiennes.

Simulation de variables réelles : fonctionnalité python

Les méthodes que nous venons de voir sont bien évidemment pré-encodées dans la plupart des langages de programmation.
En python la librairie numpy.random contient de nombreux exemples, mais il faut faire attention au lien entre les paramètres utilisés dans le programme et les paramètres utilisés en cours.

```
exponential(2) # renvoie une variable al\'eatoire de param\`etre 1/2
```

Voir aussi `https://numpy.org/doc/stable/reference/random/index.html#numpyrandom`