TP2 de Probabilités: GIS3

Benjamin Arras [*]

Ouvrir l'interface RStudio avec `rstudio`.

Exercices : Simulation par la méthode d'inversion

Rappel

Soit X une variable aléatoire à valeurs réelles telle que sa fonction de répartition F est continue. On introduit la fonction G définie par, pour tout $y \in]0,1[$

$$G(y) = \inf\{a \in \mathbb{R} : \ F(a) \geq y\}$$

On a montré (en cours et en travaux dirigés) que X a même loi que $G(U)$ où U est de loi uniforme sur $[0,1]$. Cette façon de tirer au sort suivant la loi de X est appelée méthode d'inversion. Dans les deux exercices qui suivent, on se propose donc de tester cette méthode sur deux exemples de lois.

Exercice 1
Soit Y une variable aléatoire de densité

$$f(x) = \frac{C}{(1+x)^2}\mathbb{I}_{[0,2]}(x), \quad x \in \mathbb{R}$$

1. Calculer C pour que f soit effectivement une densité de probabilité.
2. Simuler à l'aide de la méthode d'inversion un échantillon de taille 1000 de la loi de densité f.
3. Tracer un histogramme à 50 classes correspondant aux résultats de ces tirages.
4. Sur le même graphique, tracer le graphe de f et comparer les deux figures (**indication** : on utilisera la fonction "curve" correctement paramétrée).

Exercice 2
Soit Z une variable aléatoire de loi de Cauchy, c'est à dire une variable aléatoire réelle de densité

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}$$

1. Montrer que Z n'a pas de moyenne.
2. Simuler un échantillon de taille 1000 de la loi de Cauchy.

[*]Université de Lille; benjamin.arras@univ-lille.fr

3. On note $(z_1, \dots, z_{1000})$ les résultats de ces tirages. Tracer l'extrapolation linéaire de la famille de points $(n, 1/n \sum_{k=1}^{n} z_k)_{1 \leq n \leq 1000}$.

Exercices : Simulation par la méthode de rejet

Description de la méthode
Soient f et g deux densités de probabilités sur $\mathbb{R}$ pour lesquelles il existe un réel c tel que pour tout $x \in \mathbb{R}$, on a $f(x) \leq cg(x)$. On suppose qu'il est facile d'obtenir des tirages suivant la densité g et on se propose d'étudier une façon d'en tirer des tirages suivant la densité f connue sous le nom de méthode du rejet. Pour cela on introduit une suite $(Y_n)_{n\geq 1}$ de variables aléatoires indépendantes de même densité g et $(U_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de loi uniforme sur $[0,1]$. On suppose que les suites $(Y_n)_{n\geq 1}$ et $(U_n)_{n\geq 1}$ sont indépendantes. Soit h la fonction définie par

$$h(x) = \frac{f(x)}{cg(x)} \mathbb{I}_{g(x)>0}.$$

Finalement, soient X et N les variables aléatoires définies par $N = \inf\{n \geq 1,\, U_n \leq h(Y_n)\}$ et $X = Y_N$. On montrera en Probabilités 2 que la variable aléatoire X admet f pour densité. On se propose de tester cette méthode sur un exemple univarié et sur un exemple multivarié.

Exercice 3
Soit X une variable aléatoire de densité

$$f(x) = Ce^x \mathbb{I}_{[0,1]}(x), \quad x \in \mathbb{R}$$

1. Calculer C pour que f soit effectivement une densité de probabilité. Calculer $\mathbb{E}(X)$.
2. Simuler à l'aide de la méthode de rejet un échantillon de taille 1000 de la loi de densité f.
3. Tracer un histogramme à 50 classes correspondant aux résultats de ces tirages.
4. Sur le même graphique, tracer le graphe de f et comparer les deux figures.

Exercice 4
Soit $X = (X_1, X_2)$ un vecteur aléatoire de loi uniforme dans le carré $C = [-1,1]^2$.
1. Tracer sur un graphique 100 réalisations indépendantes de la loi uniforme dans le carré C.
2. On veut maintenant simuler un vecteur aléatoire $Y = (Y_1, Y_2)$ de loi uniforme dans le disque unité $D = \{(x_1, x_2) \in \mathbb{R}^2 : \sqrt{x_1^2 + x_2^2} \leq 1\}$, dont la densité est donnée par

$$f(y_1, y_2) = \frac{1}{\pi} \mathbb{I}_D(y_1, y_2) .$$

Nous allons pour cela utiliser la *méthode de rejet* adaptée au cadre de la dimension 2. Décrivons celle-ci : soit $(X_n)_{n\geq 1}$ une suite de vecteurs aléatoires indépendants et de loi uniforme dans le carré C. Notons $T := \inf\{n \geq 1 : X_n \in D\}$. Alors la v.a. T suit la loi géométrique de paramètre $\pi/4$, et X_T est un vecteur aléatoire de loi uniforme dans D.
3. Montrer l'assertion précédente.
4. Par la méthode de rejet, simuler 100 réalisations indépendantes de loi uniforme dans le disque unité D. Représenter les par un graphique. Combien de réalisations de vecteurs aléatoires de loi uniforme dans le carré C ont été nécessaires pour obtenir ces 100 réalisations (indépendantes) de loi uniforme dans D ?

5. Calculer la moyenne arithmétique de la distance séparant ces 100 réalisations de l'origine. Ce résultat est-il surprenant ?