Simulation de variables et vecteurs aléatoires

1 Générateurs Pseudo-Aléatoires.

Le logiciel Matlab, contraction de Matrix Laboratory, est un interpréteur de commandes écrites en langage Matlab. Ce logiciel dispose en particulier de certains outils utiles en probabilités et statistiques comme plusieurs générateurs de nombres aléatoires dont *rand*, associé à la uniforme sur l'intervalle $[0,1]$ et *randn* associé à la loi normale $\mathcal{N}(0,1)$. Avec la commande $rand('state', sum(100 * clock))$, le générateur *rand* est initialisé avec une graine dont la valeur dépend de l'heure. Lorsque l'on utilise *randn*, il faut de même initialiser son générateur. Les appels successifs à *rand* et *randn* fournissent des réalisations de suites de variables aléatoires indépendantes de loi $\mathcal{U}([0,1])$ et $\mathcal{N}(0,1)$, respectivement.

2 Méthode par Inversion.

Soit X une variable aléatoire réelle de fonction de répartition F. On appelle inverse généralisée de F, la fonction F^{-1} définie pour tout $y \in]0,1]$ par $F^{-1}(y) = \inf\{x \in \mathbb{R} \,/ F(x) \geq y\}$.

Lemme 1. Si U est une variable aléatoire de loi uniforme $\mathcal{U}([0,1])$, alors $F^{-1}(U)$ a même loi que X. De plus, si F est continue sur $\mathbb{R}$, alors $F(X)$ suit la loi uniforme $\mathcal{U}([0,1])$.

Exercice 1. Utiliser le code Matlab suivant pour générer N réalisations de variables aléatoires indépendantes et de loi exponentielle $\mathcal{E}(\lambda)$ et de loi de Cauchy $\mathcal{C}(c)$ avec $\lambda, c > 0$.

$N = input$('Entrez la taille de l'échantillon N : ');
$\lambda = input$('Préciser la valeur du paramètre λ : ');
$c = input$('Préciser la valeur du paramètre c : ');
$X = rand(N,1);\; Y = -\log(X)/\lambda;\; Z = c * \tan(\pi * (X - 0.5));$

Tracer les moyennes empiriques successives de Y et vérifier la loi des grands nombres pour Y. Augmenter le nombre N de réalisations pour affiner la précision. Comparer vos résultats de simulations avec le générateur *rexpweib* de Stixbox. Que se passe-t-il sur Z ? Quelle est la loi de la moyenne empirique associée à Z ? Conclure.

3 Méthode par Troncature.

Lemme 2. Si X est une variable aléatoire à valeurs dans $\mathbb{Z}$, de fonction de répartition F et soit G la fonction de répartition d'une variable aléatoire continue Y telle que, pour tout $n \in \mathbb{Z}$, $G(n+1) = F(n)$. Si U est une variable aléatoire de loi uniforme $\mathcal{U}([0,1])$, alors la partie entière $[G^{-1}(U)]$ a même loi que X.

Exercice 2. Utiliser le code Matlab suivant pour générer N réalisations de variables aléatoires indépendantes et de loi uniforme $\mathcal{U}(\{1, 2, \cdots, a\})$ avec $a \in \mathbb{N}^*$ ou bien de loi géométrique $\mathcal{G}(p)$ avec $p = 1 - \exp(-\lambda)$ et $\lambda > 0$.

$N = input($'Entrez la taille de l'échantillon N : '$);$
$a = input($'Préciser la valeur du paramètre a : '$);$
$if\ a\ \ = round(a)$
    $disp($'La valeur du paramètre a doit etre un entier positif !!'$)$
    $break$
end
$\lambda = input($'Préciser la valeur du paramètre λ : '$);$
$X = rand(N, 1);\ Y = fix(1 + a * X);\ Z = fix(-\log(X)/\lambda);$

Tracer les moyennes empiriques successives de Y et vérifier la loi des grands nombres pour Y. Augmenter le nombre N de réalisations pour affiner la précision. Effectuer le même exercice pour la loi géométrique associée à Z. Comparer vos résultats de simulations avec le générateur *rgeom* de Stixbox.

4 Lois discrètes à support fini.

Lemme 3. 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

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

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 3. Créer un code Matlab permettant de générer un vecteur aléatoire X contenant N réalisations indépendantes et de même loi Binomiale $\mathcal{B}(n, p)$ où les valeurs $N, n \geq 1$ et $0 < p < 1$ sont affectées par l'utilisateur. Pour N assez grand, vérifier la loi des grands nombres sur les moyennes empiriques successives de X. Comparer vos résultats de simulations avec le générateur *rbinom* de Stixbox.

5 Loi Normale.

Lemme 4. Soit (X, Y) un couple aléatoire de $\mathbb{R}^2$. Alors, (X, Y) suit la loi normale $\mathcal{N}(0, I_2)$ si et seulement si $X = r\cos\theta$ et $Y = r\sin\theta$ où r et θ sont deux variables aléatoires indépendantes avec r^2 de loi exponentielle $\mathcal{E}(1/2)$ et θ de loi uniforme $\mathcal{U}([0, 2\pi])$.

Il découle du lemme 4 que, 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 4. Utiliser l'algorithme de Box-Muller suivant pour générer N réalisations de variables aléatoires indépendantes et de loi normale $\mathcal{N}(m, \sigma^2)$ où la moyenne $m \in \mathbb{R}$ et la variance $\sigma^2 > 0$ sont affectées par l'utilisateur. Tracer également l'histogramme associé.

$N = input($'Entrez la taille de l'échantillon N : '$);$
$m = input($'Préciser la valeur de la moyenne m : '$);$
$\sigma^2 = input($'Préciser la valeur de la variance σ^2 : '$);$
$X = m * ones(N, 1) + sqrt(\sigma^2) * sqrt(-2 * log(rand(N, 1))). * cos(2 * pi * rand(N, 1));$
$[E, C] = histo(X, sqrt(N), 0, 1);$$hold\ on$
$title($'Simulation d'une loi normale'$);\ xlabel($'Valeurs'$);\ ylabel($'Effectifs'$);$
$plot(C, dnorm(C, m, \sigma),' r-');\ legend($'Empirique','Théorique'$);\ hold\ off$