

# TP 1 - SIMULATION DE VARIABLES ALÉATOIRES CONTINUES - AGRÉGATION EXTERNE DE MATHÉMATIQUES - 2006/2007

## 1 Algorithme de Box-Muller

1. Soient  $U$  et  $V$  deux variables aléatoires i.i.d. de loi uniforme sur  $[0, 1]$ . On rappelle que les variables aléatoires  $X$  et  $Y$  définies par

$$X = \sqrt{-2\ln(U)} \cos(2\pi V) \quad \text{et} \quad Y = \sqrt{-2\ln(U)} \sin(2\pi V)$$

sont indépendantes et de même loi normale standard  $\mathcal{N}(0, 1)$ .

2. Utilisez le résultat précédent (algorithme de Box-Muller) et la fonction `rand` pour simuler une v.a. de loi normale standard : effectuer  $N$  simulations, et comparer sur une même figure l'histogramme des valeurs obtenues et la fonction densité théorique. Que pensez-vous du résultat ?

3. Faire directement  $N$  simulations d'une loi normale standard à l'aide de la fonction `randn`. Que pensez-vous du résultat ?

## 2 Simulation d'un vecteur gaussien

1. On rappelle que si  $X = (X_1, \dots, X_k)$  est un vecteur gaussien de moyenne nulle et de covariance Identité, et  $A \in \mathcal{M}_k(\mathbb{R})$  est telle que  $AA^t = C$ , alors  $Y = AX + \mu$  est un vecteur gaussien de moyenne  $\mu$  et de covariance  $C$ . En déduire un algorithme de simulation de vecteurs gaussiens.

2. Application : faire  $N$  simulations d'un vecteur gaussien de covariance  $C$  et de moyenne  $\mu$  avec  $C = \begin{bmatrix} 5 & 3 & 2 \\ 3 & 4 & 1 \\ 2 & 1 & 2 \end{bmatrix}$  et  $\mu = [-2 \ 1 \ 4]$  (utiliser la fonction `chol`). Vérifier les résultats à l'aide des fonctions `mean` et `cov`.

## 3 Théorème de la limite centrale

1. Faire  $N$  tirages indépendants de  $n$  variables aléatoires indépendantes uniformes sur  $[-1/2, 1/2]$ . En utilisant la fonction `sum`, obtenir  $N$  tirages indépendants de la somme de  $n$  v.a. indépendantes uniformes sur  $[-1/2, 1/2]$ , normalisée par le facteur  $n^{-1/2}$  et par leur écart-type.

2. Faire l'histogramme des  $N$  valeurs obtenues, et le comparer sur un même graphique à la fonction densité de la loi normale standard  $\mathcal{N}(0, 1)$ .

3. Tracer sur un même graphique, mais avec des couleurs différentes, les fonctions de répartition empiriques (consulter `help cumsum`) obtenues pour des valeurs de  $n$  allant de 1 à 6, et comparer à la fonction de répartition de la loi normale standard (`help erf`).

## 4 Méthode d'inversion

On rappelle le résultat suivant sur lequel s'appuie la méthode d'inversion :

La fonction de répartition d'une variable aléatoire  $X$  étant définie par  $F(x) = \mathbf{P}(X \leq x)$ , si

$U$  est une variable aléatoire uniforme sur  $[0, 1]$ , alors  $X = F^{-1}(U)$  est une variable aléatoire de fonction de répartition  $F$ .

1. Utiliser cette méthode pour simuler une loi exponentielle de paramètre  $\lambda$ . Faire  $N$  simulations et comparer sur une même figure l'histogramme des valeurs obtenues avec la densité théorique.
2. Faire de même pour une loi de Cauchy.
3. Faire de même pour une loi de Weibull de paramètres  $a$  et  $b$  strictement positifs, dont la densité est donnée par

$$\forall x \geq 0, f(x) = abx^{b-1} \exp(-ax^b).$$

## 5 Loi des grands nombres

1. Faire  $N$  simulations  $X_1, \dots, X_N$  d'une loi exponentielle de paramètre  $\lambda$ , et tracer ensuite sur un même graphique la fonction  $n \mapsto (X_1 + \dots + X_n)/n$  et la valeur  $\mathbf{E}(X)$ . 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.

## 6 Non vieillissement de la loi exponentielle

1. Montrer que si  $X$  est une variable aléatoire de loi exponentielle alors, pour tout  $x, s \geq 0$ , on a  $\mathbf{P}(X \leq x + s | X \geq s) = \mathbf{P}(X \leq x)$ .
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  $X \geq s$ ".

## 7 Méthode de rejet

1. On rappelle le résultat suivant sur lequel s'appuie la méthode de rejet :  
Soient  $D$  et  $D'$  deux domaines de  $\mathbb{R}^k$  mesurables, de mesure finie et non nulle, et tels que  $D \subset D'$ . Soit  $X$  un point aléatoire uniforme sur  $D'$ , alors la loi conditionnelle de  $X$  sachant " $X \in D$ " est la loi uniforme sur  $D$ .

Cette proposition se traduit par l'algorithme suivant pour simuler une variable aléatoire uniforme sur  $D$  :

Répéter

simulation de  $X$  uniforme dans  $D'$

Jusqu'à  $(X \in D)$ .

Quelle est la loi du nombre de parcours de cette boucle ? Quelle en est sa moyenne ?

2. Application : faire  $N$  simulations de la loi uniforme sur le disque unité de  $\mathbb{R}^2$  et visualiser les points obtenus.

3. Toujours à propos de la simulation de la loi uniforme sur le disque, que pensez-vous de la méthode suivante : on commence par choisir  $x$  uniforme entre  $-1$  et  $1$  puis on prend  $y$  uniforme entre  $-\sqrt{1-x^2}$  et  $\sqrt{1-x^2}$  ? Visualiser  $N$  simulations obtenues par cette méthode. La loi est-elle uniforme ?

4. La méthode du rejet peut se généraliser à des lois non uniformes :

Soient  $f$  et  $g$  deux densités de probabilité sur  $\mathbb{R}^k$  telles qu'il existe une constante  $c$  vérifiant :  $\forall x \in \mathbb{R}^k, cg(x) \geq f(x)$ . Soit  $X$  une variable aléatoire de densité  $g$  et  $U$  une variable aléatoire de loi uniforme sur  $[0, 1]$ , indépendante de  $X$ . Soit  $E$  l'événement " $cUg(X) < f(X)$ ". Alors la loi conditionnelle de  $X$  sachant  $E$  a pour densité  $f$ .

On pourra démontrer cette proposition (on pourra commencer par calculer la probabilité de

*E*) avec les éminents professeurs du LSP en charge du cours.

L'algorithme de simulation de la loi  $f$  à partir de  $g$  est alors le suivant :

Répéter

simulation de  $X$  de densité  $g$

simulation de  $U$  uniforme sur  $[0, 1]$

Jusqu'à  $(cUg(X) < f(X))$ .

Quelle est la loi du nombre de passages dans la boucle ?

**5.** Application : on considère la densité de probabilité  $f$  définie par  $f(x) = 1 - |x|$  si  $x \in [-1, 1]$  et 0 sinon. Simuler cette loi par la méthode de rejet à partir de la loi uniforme sur  $[-1, 1]$ .