R&D

Fiche de synthèse sur la simulation

Responsable du département Statistique : Michèle DESENFANT

Fiche réalisée par Caroline Bernard-Michel

Mai 2002

- Diffusion générale –

Les techniques de simulation : méthodes générales et cas particuliers.

A. Simulation d'une loi uniforme sur [0,1]

Nous verrons par la suite que la simulation d'une variable aléatoire de loi quelconque nécessite la simulation d'une variable uniforme entre 0 et 1. D'où l'importance d'avoir un très bon générateur de loi uniforme.

Pour simuler une loi uniforme, on utilise des générateurs « pseudo-aléatoires ». Ils fournissent à l'aide d'un calcul déterministe des nombres appartenant à [0,1] ayant un comportement réputé indiscernable de ceux que fournirait une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi uniforme. Ces générateurs prennent leurs valeurs dans un ensemble fini et sont donc nécessairement périodiques. Il faut donc toujours vérifier que leur période est suffisamment grande pour ne pas avoir deux fois la même séquence de nombres dans une simulation.

Plusieurs types de générateurs «pseudo-aléatoires » sont utilisés :

- ***les générateurs congruentiels*.**

Ils sont définis par le choix d'une valeur initiale et d'une relation de récurrence :

$$I_{n+1} = (aI_n + c) \bmod b$$

- ***les générateurs de Fibonnaci*.**

Ils sont définis par le choix de deux valeurs initiales et d'une relation de récurrence :

$$I_n = (I_{n-1} + I_{n-2}) \bmod \mathrm{b}$$

- ***les générateurs de type « Multiplication avec retenue ».***

Ils sont définis par le choix de deux valeurs initiales et les relations de récurrence suivantes :

$$\begin{cases} I_{n+1} = (aI_n + c_n) \bmod \mathrm{b} \\ c_{n+1} = (aI_n + c_n) \operatorname{div} \mathrm{b} \end{cases}$$

où a est le multiplicateur, b la base, c_n la retenue.

- ***d'autres.***

B. « Méthode de la fonction de répartition » ou « méthode de la transformation inverse »

<u>RAPPELS</u> :

- <u>Définition</u> : La fonction de répartition d'une variable aléatoire X est l'application F de IR dans [0,1] définie par : $F(x) = P(X \leq x)$.
- Précisons qu'une fonction de répartition caractérise une loi, c'est à dire que si deux variables aléatoires ont la même fonction de répartition alors elles suivent la même loi.

OBJECTIF :

Soit à construire un échantillon de n réalisations d'une variable aléatoire X de fonction de répartition F

---

1 . Cas particulier où F est inversible.

- Rappelons que F et F^{-1} sont des fonctions croissantes. On a donc :

$$\forall u \in [0,1],\quad x \in \Re,\quad F^{-1}(u) \leq x \Leftrightarrow u \leq F(x) \qquad (1)$$

- Soit U une variable aléatoire de loi uniforme sur [0,1]

<u>RAPPEL</u> :La fonction de répartition G d'une loi uniforme sur [0,1] est :

$$G(x) = x \qquad \forall x \in \Re \qquad (2)$$

- D'après (1) et (2) on a :

$$\forall x \in \Re, P[F^{-1}(U) \leq x] = P[U \leq F(x)] = F(x)$$

donc la variable aléatoire $F^{-1}(U)$ suit une loi de fonction de répartition F.

$\Rightarrow$ X et $F^{-1}(U)$ suivent la même loi.
$\Rightarrow$ Pour simuler X, il faut simuler $F^{-1}(U)$.

RESUME METHODE

Pour générer un échantillon $x_{1,} x_{2,} ...., x_n$ d'une variable aléatoire X de loi quelconque de fonction de répartition F, il faut :

1. Si F est explicite et inversible, calculer F^{-1}
2. Générer un échantillon $u_1, u_2, ..., u_n$ d'une variable aléatoire U uniformément distribuée sur [0,1] et poser X=F^{-1}(U)

EXEMPLE

On veut simuler une loi exponentielle de paramètre λ de densité $f(x) = \lambda e^{-\lambda x} 1_{IR^+}$

1. Sa fonction de répartition est : $F(X) = P(X \leq x) = \int_0^x \lambda e^{-\lambda x} \, dx = 1 - e^{-\lambda x}$
2. On calcule l'inverse de F : $F^{-1}(u) = \dfrac{-\ln(1-u)}{\lambda}$ $\qquad u \in [0,1[$
3. Soit U une variable aléatoire de loi uniforme sur [0,1].

On génère n réalisations de U à partir d'un des générateurs présentés en début de document, puis on pose $X = \dfrac{-\ln(1-U)}{\lambda}$. On obtient un échantillon de n valeurs de X.

2. Cas où F n'est pas inversible

Au lieu d'utiliser F^{-1}, on utilise la fonction de quantiles q sur [0,1] :

$$q(u) = \inf\{y \in IR, F(y) \geq u\}.$$

$\Rightarrow$ Pour simuler X il suffira de simuler q(U)

Remarques :

- la démonsration est la même que celle du cas numéro 1. On utilise juste en plus la continuité à droite de F.
- Si F est inversible alors $q = F^{-1}$

C. Cas de la loi normale

1. Par la méthode générale

La fonction de répartition d'une loi normale n'est pas explicite. On peut utiliser des approximations de cette fonction mais on risque d'accumuler les erreurs.

2. Méthode du théorème central limite.

Soit Y une variable aléatoire uniforme sur [0,1] alors :

$E(Y)=\frac{1}{2}$ et $VAR(Y)=\frac{1}{12}$.

Soient $Y_1, Y_2, ...Y_m$ variables aléatoires uniformes sur [0,1]

D'après le théorème central limite, la variable définie par : $\dfrac{\sum_{i=1}^{m} Y_i - \frac{m}{2}}{\sqrt{\frac{m}{12}}}$ tend vers une loi normale centrée réduite N(0,1) quand m est grand.

RESUME METHODE :

Pour générer un échantillon $x_1, x_2, ...x_n$ d'une variable aléatoire X de loi normale il faut :

- Générer m échantillons de taille n des m variables $U_1, U, ...U_m$ uniformément distribuées sur [0,1] :

$$
\begin{array}{l}
(u_{11}, u_{12}, ...u_{1n}) \\
.... \\
.... \\
(u_{m1}, u_{m2}, ...u_{mm})
\end{array}
$$

- Poser X= $\dfrac{\sum_{i=1}^{m} U_i - \frac{m}{2}}{\sqrt{\frac{m}{12}}}$, on a donc $x_j = \dfrac{\sum_{i=1}^{m} u_{ij} - \frac{m}{2}}{\sqrt{\frac{m}{12}}}$

3. Méthode de Box Muller

On utilise le théorème suivant :

Si U_1 et U_2 sont indépendants et de loi uniforme sur [0,1], les variables aléatoires suivantes :

- $X_1 = (-2\log U_1)^{1/2}\cos(2\pi U_2)$
- $X_2 = (-2\log U_1)^{1/2}\sin(2\pi U_2)$

sont deux variables gaussiennes indépendantes centrées réduites N(0,1).

RESUME METHODE:

Pour générer un échantillon $x_1, x_2, ...x_n$ d'une variable aléatoire X de loi normale, il faut :

- générer deux échantillons ($u_1, u_2, ...u_n$) et ($v_1, v_2, ...v_n$) de deux variables U et V uniformément distribuées sur [0,1] et indépendantes.
- poser $X_1 = (-2\log U_1)^{1/2}\cos(2\pi U_2)$ ou $X_2 = (-2\log U_1)^{1/2}\sin(2\pi U_2)$

Remarque :

Pour simuler une variable aléatoire N(m, σ^2) il suffit de poser $X = m + \sigma G$ ou G suit une loi normale centrée réduite.