

IFT 3245

Simulation et modèles

Fabian Bastin
DIRO
Université de Montréal

Automne 2016

Il s'agit de la technique la plus importante après l'inversion, et est parfois (mais pas toujours !) compatible avec l'utilisation de techniques de réduction de variance. Contrairement à la technique d'inversion, elle ne requiert pas de pouvoir facilement calculer la fonction de répartition, et encore moins son inverse.

Nous allons considérer le cas où X est continu (le cas discret est analogue).

Fonction majorante

Soit $f(x)$ la densité de X , et soit t une fonction majorant f , i.e. $f(x) \leq t(x)$ pour tout x . On peut normaliser t comme

$$r(x) = t(x)/a, \text{ où } a = \int_{-\infty}^{\infty} t(s)ds.$$

r est une densité comme son intégration donne 1 :

$$\int_{-\infty}^{\infty} r(x) = \int_{-\infty}^{\infty} t(s)/ads = \frac{1}{a} \int_{-\infty}^{\infty} t(s)ds = 1.$$

On choisit t de manière à ce que

- (i) ce soit facile de générer des v.a. de densité r ,
- (ii) a soit petit (proche de 1), ou en d'autres termes $t(x)$ est proche de $f(x)$.

Le choix de t peut être automatisé.

Méthode d'acceptation-rejet

Répéter

- Générer Y de densité $r(x)$;
- Générer $U : U(0, 1)$ indépendante de Y ;

jusqu'à obtenir $U \leq f(Y)/t(Y)$; retourner Y .

La variable aléatoire Y retournée est de densité f .

Convaincu ?

Démonstration.

Nous avons tout d'abord, comme $P[U \leq x] = x$,

$$\begin{aligned}P[Y \text{ est accepté}] &= P[U \leq f(Y)/t(Y)] \\&= \int_{-\infty}^{\infty} P[U \leq f(y)/t(y)]r(y)dy \\&= \int_{-\infty}^{\infty} f(y)/t(y)r(y)dy \\&= \frac{1}{a}.\end{aligned}$$



Démonstration.

Similairement,

$$\begin{aligned} & P[Y \leq x \text{ et } Y \text{ est acceptée}] \\ &= \int_{-\infty}^{\infty} P[Y \leq x \text{ et } U \leq f(Y)/t(Y) | Y = y] r(y) dy \\ &= \int_{-\infty}^{\infty} P[y \leq x \text{ et } U \leq f(y)/t(y)] r(y) dy \\ &= \int_{-\infty}^x P[U \leq f(y)/t(y)] r(y) dy \\ &= \int_{-\infty}^x [f(y)/t(y)] r(y) dy \\ &= F(x)/a \end{aligned}$$



Démonstration.

Nous pouvons dès lors écrire

$$\begin{aligned} P[Y \leq x | Y \text{ est accepté}] &= \frac{P[Y \leq x \text{ et } Y \text{ est accepté}]}{P[Y \text{ est accepté}]} \\ &= \frac{F(x)/a}{1/a} = F(x). \end{aligned}$$



Méthode d'acceptation-rejet

À chaque tour de boucle, la probabilité d'accepter Y est $1/a$.

Le nombre de tours de boucle avant l'acceptation est une v.a. géométrique de paramètre $p = 1/a$.

Le nombre moyen de tours de boucle par v.a. est donc a .

Méthode d'acceptation-rejet : exemple

Supposons que nous souhaitons tirer des valeurs d'une loi beta de paramètres (4,3), de fonction de densité

$$f(x) = 60x^3(1 - x)^2, \text{ pour } 0 < x < 1.$$

$F(x)$ est difficile à inverser, mais il est facile d'établir que le maximum de f est $f(0.6) = 2.0736$.

On peut donc prendre $t(x) = 2.0736$ pour $0 < x < 1$. Dans ce cas $r(x)$ est la densité $U(0, 1)$.

On peut diminuer a en choisissant une densité t un peu moins simpliste.

Bien que les méthodes décrites dans les sections précédentes décrivent les techniques majeures pour générer des variables aléatoires multivariées, il existe de nombreuses approches spécifiques aux distributions considérées.

Ces techniques sont cependant rarement compatibles avec les techniques d'amélioration d'efficacité, et dès lors devraient être employées parcimonieusement.

Référence utile : Luc Devroye, *Non-Uniform Random Variate Generation*, <http://www.nrbook.com/devroye/>.

Méthode de Box-Muller pour la loi normale

Génération de variables aléatoires normales en suivant une transformée de coordonnées cartésiennes en coordonnées polaires.

La méthode proposée en 1958 par Box et Muller ; exacte, mais plus lente que l'inversion.

Son intérêt réside donc d'avantage au niveau théorique qu'au niveau pratique, mais ce type de transformation peut se révéler utile dans d'autres cadres.

Méthode de Box-Muller pour la loi normale

Considérons un couple (X, Y) de variables aléatoires $N(0, 1)$ indépendantes. Leur densité jointe sur \mathcal{R}^2 est

$$f(x, y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}.$$

Il est aisé de changer les coordonnées cartésiennes (X, Y) par les coordonnées polaires (R, Θ) :

$$R^2 = X^2 + Y^2; Y = R \sin \Theta.$$

La transformée inverse est donnée par $X = R \cos \Theta$ et $Y = R \sin \Theta$. La matrice jacobienne de la transformation (r, θ) vers (x, y) est

$$J = \begin{pmatrix} \cos \theta & -R \sin \theta \\ \sin \theta & R \cos \theta \end{pmatrix},$$

de déterminant

$$(\cos \theta)(r \cos \theta) - (\sin \theta)(-r \sin \theta) = r(\cos^2 \theta + \sin^2 \theta) = r.$$

Méthode de Box-Muller pour la loi normale

Par conséquent, (R, Θ) a la densité

$$t(r, \theta) = |J|f(x, y) = rf(r \cos \theta, r \sin \theta) = (r/2\pi)e^{-r^2/2}.$$

Puisque cette densité ne dépend pas de θ , Θ doit avoir la densité uniforme sur $[0, 2\pi]$. En intégrant par rapport à θ sur l'intervalle $[0, 2\pi]$, nous obtenons que R a la densité

$$f_R(r) = \int_0^{2\pi} t(r, \theta) d\theta = re^{-r^2/2} \quad \text{pour } r \geq 0.$$

La fonction de répartition correspondante est

$F_R(r) = 1 - e^{-r^2/2}$, et son inverse est donnée par

$$F_R^{-1}(U) = \sqrt{-2 \ln(1 - U)}.$$

Méthode de Box-Muller pour la loi normale

Il suffit alors de générer θ et R indépendamment, en utilisant U_1 et U_2 , puis à transformer ces coordonnées polaires en coordonnées rectangulaires (X, Y) , autrement dit

$$X = R \cos \theta = \cos(2\pi U_1) \sqrt{-2 \ln(U_2)}$$

$$Y = R \sin \theta = \sin(2\pi U_1) \sqrt{-2 \ln(U_2)}.$$

Uniforme sur la sphère unité

Nous souhaitons générer un vecteur U uniformément distribué sur $\mathcal{C}^k = \{x \in \mathcal{R}^k \mid \|x\|_2 = 1\}$. Ce genre de variables apparaît dans la génération de variables multivariées elliptiques, utilisées notamment en gestions de risques. La génération de fait en 3 lignes de codes !

```
void generate_hypersphere
(Random *ran, int k, double *x)
{
    int i;
    double norm;

    for (i = 0; i < k; i++)
        x[i] = ran_normal_icdf(ran, 0, 1);
    norm = cblas_dnrm2(n, x, 1);
    cblas_dscal(n, 1/norm, x, 1);
}
```

Uniforme sur la sphère unité

(Devroye, Chap. 5, Section 4)

Si N_1, \dots, N_k sont des normales i.i.d., alors, en notant $N = (N_1, \dots, N_k)$, $U = N/\|N\|_2$ est uniformément distribué sur \mathcal{C}^k .

Si $k = 2$, nous pouvons calculer un tel point en définissant une variable aléatoire Θ uniformément distribuée sur $[0, 2\pi)$.

Toutefois, le raisonnement ne peut pas s'étendre pour $k \geq 3$.

Considérons $k = 3$. Nous serions alors tentés d'utiliser deux variables aléatoires uniformément distribuées sur $[0, 2\pi)$. Notons ces variables Θ et B , et supposons les indépendantes.

Uniforme sur la sphère unité

Nous pouvons faire correspondre à chaque intervalle de $[0, 2\pi)$ une probabilité, dont la valeur est fixée par la différence entre les extrémités de l'intervalle. Considérons la surface décrite en laissant varier la première variable aléatoire Θ de θ_1 à θ_2 , et la seconde, B , de β_1 à β_2 . La surface vaut, en prenant R comme rayon de la sphère,

$$\begin{aligned} S &= \left| \int_{\beta_1}^{\beta_2} \int_{\alpha_1}^{\alpha_2} R^2 \cos \beta d\alpha d\beta \right| \\ &= \left| R^2 (\alpha_2 - \alpha_1) (\sin \beta_1 - \sin \beta_2) \right|. \end{aligned}$$

Par analogie avec la cartographie, nous dénommerons α la longitude et β la latitude.

Uniforme sur la sphère unité

Si la longueur de l'arc décrit en fonction de la latitude ne dépend que de la différence entre les latitudes extrême, il n'en va pas de même pour la longitude : l'arc décrit est d'autant plus petit que la latitude est importante.

Si les points étaient uniformément distribués sur la sphère, à deux surfaces de même valeur devraient correspondre à des probabilités identiques. Nous pouvons constater que suivant la latitude, nous devons varier la différence dans les longitudes.

De manière équivalente, nous pouvons dire que les surfaces ne dépendent pas uniquement des différences d'angles. Par conséquent, les points ne sont pas distribués uniformément sur la sphère.