UNIVERSITÉ DE NICE SOPHIA-ANTIPOLIS
Master 1 IMEA - Méthodes de Monte-Carlo
Sylvain Rubenthaler
http ://math.unice.fr/∼rubentha/cours.html

Feuille d'exercices numéro 9

1. Soit X v.a.r. de fonction de répartition F que l'on supposera inversible.
   (a) Comment simuler X conditionellement à $X > m$ à l'aide d'une méthode de rejet ? Évaluer l'efficacité de la méthode ? Que se passe-t-il quand m devient grand ?
   (b) Soit $U \sim \mathcal{U}([0,1])$ et $Z = F^{-1}(F(m) + (1 - F(m))U)$. Montrer que $Z \sim \mathcal{L}(X|X > m)$. Comparer cette méthode de simulation avec la précédente.
   (c) Comment simuler suivant $\mathcal{L}(X|a < X < b)$?
   (d) Supposons maintenant que l'on cherche à simuler suivant $\mathcal{L}(X|X > m)$ pour $X \sim \mathcal{N}(\mu, \sigma^2)$. Montrer que l'on peut se rammener au cas $\mu = 0, \sigma = 1$.
   (e) Proposer une méthode de rejet (pour simuler suivant $\mathcal{L}(X|X > m)$) basée sur la loi de densité $\theta e^{-\theta(x-m)}\mathbf{1}_{x>m}$. Comment choisir θ de manière optimale ?
2. On veut évaluer $I = \int_{\mathbb{R}} f(x)p(x)dx$ par une méthode de Monte-Carlo en simulant un échantillon $X_1, \dots, X_n$ de densité p par la méthode du rejet. On a une densité q vérifiant $p \leq Mq$. On simule des variables Z_i suivant q, on note $X_1, ..., X_n$ celles que l'on garde (et qui ont donc une loi p) et $Z_1, ..., Z_{N-n}$ celles que l'on rejette. Le nombre n est fixé et N est le nombre de variables qu'on a dû simuler pour récupérer n variables de loi de densité p.
   (a) Loi de N ?
   (b) Loi de $(Z_1, \dots, Z_{N-n})$ conditionellement à $N - n = k$?
   (c) Montrez que
   $$\frac{1}{N}\left(\sum_{i=1}^{n} f(X_i) + \sum_{i=1}^{N-n} \frac{(M-1)p(Z_i)}{Mq(Z_i) - p(Z_i)} f(Z_i)\right)$$
   est un estimateur sans biais de $\mathbb{E}(f(X_1))$ (c'est à dire que son espérance est $\mathbb{E}(f(X_1))$).
3. Soit $N \in \mathbb{N}^*$. Soient $\xi_1, \dots, \xi_N$ dans $\mathbb{R}^d$ ($d \in \mathbb{N}^*$), $\omega_1, \dots, \omega_N \in \mathbb{R}^+$ tels que $\omega_1 + \dots + \omega_N = 1$. On aimerait tirer N variables i.i.d. suivant $\mu = \sum_{i=1}^{N} \omega_i \delta_{\xi_i}$
   (a) Soient $U_1, \dots, U_N$ i.i.d. de loi $\mathcal{U}([0;1])$. Soient des variables $Y_1, \dots, Y_N$ définies par
   $$Y_j = \xi_k \text{ si } \omega_1 + \dots + \omega_{k-1} \leq U_j < \omega_1 + \dots + \omega_k \,.$$
   Montrer que $Y_1, \dots, Y_N$ sont i.i.d. de loi μ. Proposer un algorithme pour simuler les $Y_1, \dots, Y_N$ de coût $O(N^2)$.
   (b) Soit $N \in \mathbb{N}^*$. Soient $X_1, \dots, X_{N+1}$ i.i.d. de loi $\mathcal{E}(1)$. Soit $Z = X_1 + \dots + X_{N+1}$. Soient $U_1, \dots, U_N$ i.i.d. de loi $\mathcal{U}([0;1])$. Soient $V_1 \leq \dots \leq V_N$ tels que $\{V_1, \dots, V_N\} = \{U_1, \dots, U_N\}$ (c'est la statistique d'ordre de $(U_1, \dots, U_N)$). Montrer que $Y_1 = \frac{X_1}{Z}$, $Y_2 = \frac{X_1+X_2}{Z}$, ..., $Y_N = \frac{X_1+\dots+X_N}{Z}$ sont tels que
   $$(Y_1, \dots, Y_N) \overset{\text{(LOI)}}{=} (V_1, \dots, V_N) \,.$$
   (c) Soient $\xi_1, \dots, \xi_N$ dans $\mathbb{R}^d$ ($d \in \mathbb{N}^*$), $\omega_1, \dots, \omega_N \in \mathbb{R}^+$ tels que $\omega_1 + \dots + \omega_N = 1$. Trouver une méthode pour tirer N variables i.i.d. suivant la loi μ de coût $O(N)$.

(d) (Programmation) Soit $(X_k)_{k\geq 0}$ chaîne de Markov dans $\mathbb{R}$ de loi initale δ_0 et de transition donnée par $X_{n+1} = 0,8 \times X_n + V_n$ (avec des V_n i.i.d. de loi $\mathcal{E}(1)$). Soient (pour $n \geq 1$) : $Y_n = X_n + W_n$ (avec des W_n i.i.d., indépendants de tout le reste et de loi $\mathcal{N}(0,1)$. Écrire un programme qui approche $\mathbb{E}(X_n|Y_1,\dots,Y_n)$ pour n dans un intervalle.