Université de Strasbourg — Ségolen Geffray
M1 - Magistère — geffray@math.unistra.fr
Statistique - études de cas — Année 2014/2015

Sujet 5: variations autour de la méthode de Monte-Carlo

• **Méthode des variables de contrôle**:
Le principe de la méthode consiste à trouver une variable aléatoire Y et une constante c telles que
$$\mathbb{E}[\varphi(X)] = \mathbb{E}[Y] + c \text{ et } \mathrm{Var}(Y) < \mathrm{Var}(\varphi(X))\,.$$
Cette méthode permet certes de diminuer la variabilité des entrées... mais entraîne une augmentation du temps de calcul. Pour trouver Y, on cherche une décomposition de $\mathbb{E}[\varphi(X)]$ de la forme
$$\mathbb{E}[\varphi(X)] = \mathbb{E}[\underbrace{\varphi(X) - h(X)}_{=Y}] + \underbrace{\mathbb{E}[h(X)]}_{=c}$$
où $\mathbb{E}[h(X)]$ peut être calculé analytiquement et où $\mathrm{Var}(\varphi(X) - h(X)) < \mathrm{Var}(\varphi(X))$.

• **Echantillonnage stratifié**:
On veut calculer $I = \mathbb{E}[\varphi(X)] = \displaystyle\int \varphi(x)f(x)dx$ où X est une variable aléatoire de $\mathbb{R}^d$ admettant la densité f par rapport à la mesure de Lebesgue. On introduit une partition $(A_j)_{j=1,\ldots,J}$ de $X(\Omega) \subset \mathbb{R}^d$ et on pose $p_j = \mathbb{E}[\varphi(X)|X \in A_j]$ pour $j = 1, ..., J$. On réécrit alors I sous la forme:
$$I = \sum_{j=1}^{J} \mathbb{E}[\varphi(X)|X \in A_j]\mathbb{P}(X \in A_j) = \sum_{j=1}^{J} p_j I_j\,.$$
L'idée de la méthode consiste à estimer I_j par la moyenne empirique $\widehat{I}_j$ d'un échantillon de taille n_j de la loi conditionnelle de X sachant $\{X \in A_j\}$, ce qui mène à l'estimateur suivant de I:
$$\widehat{I}_n^S = \sum_{j=1}^{J} p_j \widehat{I}_j\,.$$
L'estimateur $\widehat{I}_n^S$ est sans biais pour I et fortement consistant lorsque $n_j \to \infty$ pour $j = 1, ..., J$. Notons $\sigma_j^2 = \mathrm{Var}(\varphi(X)|X \in A_j)$. Si les estimateurs $\widehat{I}_j$ pour $j = 1, ..., J$ sont indépendants, alors la variance de $\widehat{I}_n^S$ est donnée par
$$\mathrm{Var}\left(\widehat{I}_n^S\right) = \sum_{j=1}^{J} \frac{p_j^2\sigma_j^2}{n_j}\,.$$
Si l'on dispose d'un total de $n = \displaystyle\sum_{j=1}^{J} n_j$ simulations, alors alors le minimum de $\mathrm{Var}\left(\widehat{I}_n^S\right)$ sous la contrainte de $n = \displaystyle\sum_{j=1}^{J} n_j$ est réalisé lorsque $\sigma_j = cp_j\sigma_j$ pour $j = 1, ..., J$ soit
$$n_j = n\,\frac{p_j\sigma_j}{\sum_{k=1}^{J} p_k\sigma_k}$$

et vaut

$$\operatorname{Var}\left(\widehat{I}_n^S\right) = \frac{1}{n}\left(\sum_{j=1}^{J} p_j \sigma_j\right)^2 .$$

Il reste à comparer cette variance à la variance de l'estimateur par la méthode Monte-Carlo usuelle $\widehat{I}_n = \frac{1}{n}\sum_{i=1}^{n}\varphi(X_i)$ laquelle vaut

$$\operatorname{Var}\left(\widehat{I}_n\right) = \frac{\operatorname{Var}(\varphi(X))}{n} .$$

Décomposons

$$\begin{aligned}
\operatorname{Var}(\varphi(X)) &= \mathbb{E}\left[\varphi(X)^2\right] - \mathbb{E}\left[\varphi(X)\right]^2 \\
&= \sum_{j=1}^{J} p_j \mathbb{E}\left[\varphi(X)^2 | X \in A_j\right] - \left(\sum_{j=1}^{J} p_j \mathbb{E}\left[\varphi(X) | X \in A_j\right]\right)^2 \\
&= \sum_{j=1}^{J} p_j \sigma_j^2 + \sum_{j=1}^{J} p_j \mathbb{E}\left[\varphi(X) | X \in A_j\right]^2 - \left(\sum_{j=1}^{J} p_j \mathbb{E}\left[\varphi(X) | X \in A_j\right]\right)^2
\end{aligned}$$

Utilisons la convexité de la fonction $x \to x^2$ et le fait que $\sum_{j=1}^{J} p_j = 1$ pour obtenir que

$$\sum_{j=1}^{J} p_j \mathbb{E}\left[\varphi(X) | X \in A_j\right]^2 \geq \left(\sum_{j=1}^{J} p_j \mathbb{E}\left[\varphi(X) | X \in A_j\right]\right)^2$$

et donc

$$\operatorname{Var}(\varphi(X)) \geq \sum_{j=1}^{J} p_j \sigma_j^2 .$$

En réutilisant la convexité de la fonction $x \to x^2$ et le fait que $\sum_{j=1}^{J} p_j = 1$, on en conclut que

$$\operatorname{Var}(\widehat{I}_n) \geq \frac{1}{n}\left(\sum_{j=1}^{J} p_j \sigma_j\right)^2 = \operatorname{Var}\left(\widehat{I}_n^S\right) .$$

Notons toutefois que le choix optimal de n_j demande de connaître les variances conditionnelles σ_j^2 ce qui n'est pas toujours le cas. On peut penser les estimer à leur tour par méthode de Monte-Carlo ... mais il faut prendre garde au fait qu'un mauvais choix des n_j peut augmenter la variance. Cependant, notons que le choix $n_j = np_j$ diminue toujours la variance même s'il n'est pas optimal. En effet, dans ce cas, on a

$$\operatorname{Var}\left(\widehat{I}_n^S\right) = \frac{1}{n}\sum_{j=1}^{J} p_j \sigma_j^2 \leq \frac{\operatorname{Var}\left(\varphi(X)\right)}{n} = \operatorname{Var}(\widehat{I}_n) .$$

Pour simuler des variables selon une loi uniforme stratifiée, on peut procéder comme suit:

- on choisit une partition de $]0,1] = \bigcup_{j=1}^{J} A_j$ où les A_j sont disjoints et de la forme $A_j =]a_j, b_j]$ (typiquement, on peut choisir $A_j = \left]\frac{j-1}{J}, \frac{j}{J}\right]$)
- la loi de $U \sim \mathcal{U}(0,1)$ conditionnelle à $\{U \in A_j\}$ est la loi $\mathcal{U}(a_j, b_j)$
- on génère donc un échantillon de taille n_j de $V_{i,j} \sim \mathcal{U}(a_j, b_j)$
- on fournit les variables $V_{i,j}$

Pour simuler des variables selon une loi quelconque de fonction de répartition associée notée F stratifiée, on peut procéder comme suit:

- on choisit la partition de $]0,1] = \bigcup_{j=1}^{J} A_j$ où les A_j sont de la forme $A_j =]a_{j-1}, a_j]$ avec $a_0 = -\infty$, $a_j = F^{-1}\left(\sum_{\ell=1}^{j} p_\ell\right)$ pour $j = 1, ..., J$ où l'on se donne les p_j sont tels que $0 < p_j < 1$ et $\sum_{j=1}^{J} p_j = 1$.
  NB: dans le cas où $a_J = +\infty$, on prend $A_j =]a_{J-1}, a_J[$
- la loi de $U \sim \mathcal{U}(0,1)$ conditionnelle à $\{U \in A_j\}$ est la loi $\mathcal{U}(a_j, b_j)$
- on génère un échantillon de taille n_j de $V_{i,j} \sim \mathcal{U}(a_{j-1}, a_j)$
- on fournit les variables $Y_{i,j} = F^{-1}(V_{i,j})$ qui suivent la loi F sachant que $Y_{i,j} \in A_j$.

Exercice 1.
Calculer une estimation de π en utilisant le fait que l'intégrale suivante vaut π:

$$I = 4\int_0^1 \sqrt{1-x^2}dx.$$

Pour cela, vous utiliserez tour à tour:

1. la méthode de Monte-Carlo usuelle,
2. la méthode des variables antithétiques,
3. la méthode de stratification,
4. la méthode de la variable de contrôle avec la variable de contrôle $Y = c\left(X - \frac{1}{2}\right)$ où $X \sim \mathcal{U}(0,1)$, et pour une constante c bien choisie.

Comparer l'efficacité des différentes méthodes mises en oeuvre.