Université de Strasbourg
M1 - Magistère
Statistique - études de cas

Ségolen Geffray
geffray@math.unistra.fr
Année 2014/2015

**Sujet 6: variations autour de la méthode de Monte-Carlo**

• **Conditionnement**:
Considérons les égalités suivantes, valable pour toutes variables $X$ et $Y$ telles que les moments mentionnés existent:

$$\begin{aligned} &\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]\,, \\ &\mathrm{Var}(X) = \mathrm{Var}\left(\mathbb{E}[X|Y]\right) + \mathbb{E}\left[\mathrm{Var}(X|Y)\right] \qquad \Rightarrow \qquad \mathrm{Var}(X) \geq \mathrm{Var}\left(\mathbb{E}[X|Y]\right)\,. \end{aligned}$$

Il s'agit donc de trouver une variable aléatoire $Y$ telle que $\mathbb{E}[X|Y] = \phi(Y)$ avec $\phi(y) = \mathbb{E}[X|Y = y]$ exploitable ie d'expression explicite.
Déjà remarquons que si $X = f(Y,Z)$ avec $Y$ et $Z$ indépendantes, alors

$$\mathbb{E}[f(Y,Z)|Y = y] = \mathbb{E}[f(y,Z)].$$

Supposons par exemple que l'on veuille calculer $\mathbb{P}(X \leq Y)$ pour deux variables aléatoires $X$ et $Y$ indépendantes, de fonctions de répartition respectives $F$ et $G$. Remarquons que

$$\mathbb{P}(X \leq Y) = \mathbb{E}[I(X \leq Y)] = \int\int I(x \leq y)dF(x)dG(y) = \int F(y)dG(y) = \mathbb{E}[F(Y)]\,.$$

Ceci n'est intéressant que si l'on sait exprimer analytiquement $F$... mais la réduction de variance peut être importante quand la probabilité $\mathbb{P}(X \leq Y)$ est petite.

• **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 $n = \sum_{j=1}^{J} n_j$ simulations, alors alors le minimum de Var $\left(\widehat{I}_n^S\right)$ sous la contrainte de $n = \sum_{j=1}^{J} n_j$ est réalisé lorsque $\sigma_j = c p_j \sigma_j$ pour $j = 1, ..., J$ soit

$$n_j = n \, \frac{p_j \sigma_j}{\sum_{k=1}^{J} p_k \sigma_k}$$

et vaut

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

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

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

Décomposons

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

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

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

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

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

Notons toutefois que le choix optimal de $n_j$ demande de connaître les variances conditionnelles $\sigma_j^2$ 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

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

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 $\pi$ en utilisant le fait que l'intégrale suivante vaut $\pi$:

$$I_3 = 4\int_0^1\int_0^1 I(x^2+y^2 \leq 1)dxdy.$$

Pour cela, vous utiliserez tour à tour:

1. la méthode de Monte-Carlo usuelle,
2. la méthode du conditionnement,
3. la méthode des variables antithétiques,
4. la méthode de la stratification.

Comparer l'efficacité des différentes méthodes mises en oeuvre.