Travaux Dirigés n°4

stoehr@ceremade.dauphine.fr

Exercice 1 *(Variable de contrôle).* On définit $\theta = \mathbb{E}[h(\mathbf{X})]$ et $m = \mathbb{E}[h_0(\mathbf{X})]$, avec $\mathbf{X} \sim \nu$. On suppose que m est connu et que l'on souhaite estimer $\delta = \theta - m$. Soit $\mathbf{X}_1, \ldots, \mathbf{X}_n$ un n-échantillon suivant ν, on considère

$$\widehat{\delta}_n^{(1)} = \frac{1}{n}\sum_{k=1}^{n} h(\mathbf{X}_i) - m \quad \text{et} \quad \widehat{\delta}_n^{(2)} = \frac{1}{n}\sum_{k=1}^{n} h(\mathbf{X}_i) - h_0(\mathbf{X}_i).$$

Pour quelle valeur de la corrélation ρ entre $h(\mathbf{X})$ et $h_0(\mathbf{X})$, l'estimateur $\widehat{\delta}_n^{(2)}$ est-il plus efficace en terme de variance que $\widehat{\delta}_n^{(1)}$? On supposera pour simplifier que $\mathbb{V}\mathrm{ar}[h(\mathbf{X})] = \mathbb{V}\mathrm{ar}[h_0(\mathbf{X})]$ est finie, que le coût d'évaluation de h et h_0 est le même et que le coût de simulation suivant ν est négligeable.

Exercice 2 *(Variables de contrôle).* On suppose que $X_1, \ldots, X_d$ sont *i.i.d.* suivant la loi de Gumbel dont la densité est donnée par

$$f(x) = \frac{1}{\beta}\exp\left\{-\exp\left(-\frac{x-\mu}{\beta}\right)\right\}\exp\left(-\frac{x-\mu}{\beta}\right)\mathbb{1}_{\{x\in\mathbb{R}\}}, \quad \text{avec} \quad \mu \in \mathbb{R} \quad \text{et} \quad \beta \in \mathbb{R}_+^*.$$

On cherche à estimer, pour $t \in \mathbb{R}_+$,

$$p = \mathbb{P}\left[\sum_{i=1}^{d} \exp(X_i) > t\right].$$

Pour $1 \le k \le n$, on notera $X_{1,k}, \ldots, X_{d,k}$, $1 \le k \le n$, n réalisations de $X_1, \ldots, X_d$. On considère les fonctions

$$h_{0,1} : x_1, \ldots, x_n \mapsto \frac{\partial}{\partial \mu}\ell(\mu, \beta \mid x_1, \ldots, x_d) \quad \text{et} \quad h_{0,2} : x_1, \ldots, x_n \mapsto \frac{\partial}{\partial \beta}\ell(\mu, \beta \mid x_1, \ldots, x_d),$$

où $\ell(\mu, \beta \mid x_1, \ldots, x_d)$ est la log-vraisemblance associée à $X_1, \ldots, X_d$ variables aléatoires *i.i.d.* de loi de Gumbel.

1. Montrer que sous des conditions de régularité usuelles $\mathbb{E}\left[h_{0,1}(X_1, \ldots, X_d)\right] = \mathbb{E}\left[h_{0,2}(X_1, \ldots, X_d)\right] = 0$.

2. En déduire deux estimateurs possibles de la variable de contrôle.

3. En pratique, quel critère utiliseriez-vous pour choisir lequel de ces estimateurs utiliser ? Vous justifierez le critère utilisé.

Exercice 3 *(Estimateur stratifié).* Le nombre de précipitation S sur un mois est modélisé par une loi de Poisson de paramètre $\lambda = 3.7$. La quantité d'eau Q_s tombant lors d'une précipitation s est modélisé par un loi de Weibul de paramètre de forme $k = 0.5$ et de paramètre d'échelle $\lambda = 2$ (on supposera que les précipitations sont indépendantes). La quantité de pluie tombant en 1 mois est donc

$$X = \begin{cases} 0 & \text{, si } S = 0, \\ \sum_{s=1}^{S} Q_s & \text{, sinon.} \end{cases}$$

On s'intéresse aux mois présentant de faibles précipitations et on cherche à estimer $p = \mathbb{P}[X < 3]$ (*i.e.*, il y a moins de 3 cm de pluie par mois).

1. Proposer un estimateur de p par la méthode de Monte Carlo classique.

2. Construire un estimateur de p par la méthode de stratification avec allocation proportionnelle, en précisant les choix faits.