Gestion des incertitudes
TD3 : Analyse de risque et évènements rares

Lucie Le Briquer

26 janvier 2018

Exercice 1 (grandes déviations et échantillonnage d'importance)

1. (X_i) i.i.d. $\tilde{X}_n = \sum_{i=1}^n X_i$. Pour $x > \mathbb{E}(X_1)$,

$$p_n(x) = \mathbb{P}\left(\frac{\tilde{X}_n}{n} \geqslant x\right) = \mathbb{E}\left(\mathbb{1}_{\frac{\tilde{X}_n}{n} \geqslant x}\right) \xrightarrow[n\to+\infty]{} 0$$

par la loi forte des grands nombres.

$$\begin{aligned} p_n\left(\mathbb{E}(X_1) + \frac{y}{\sqrt{n}}\right) &= \mathbb{P}\left(\frac{\tilde{X}_n}{n} \geqslant \mathbb{E}(X_1) + \frac{y}{\sqrt{n}}\right) \\ &= \mathbb{E}\left(\mathbb{1}_{\sqrt{n}\left(\frac{\tilde{X}_n}{n} - \mathbb{E}(X_1)\right) \geqslant y}\right) \end{aligned}$$

Or :

$$\sqrt{n}\left(\frac{\tilde{X}_n}{n} - \mathbb{E}(X_1)\right) \xrightarrow[n\to+\infty]{(\mathcal{L})} \mathcal{N}(0, \mathrm{Var}(X_1))$$

Ainsi,

$$p_n\left(\mathbb{E}(X_1) + \frac{y}{\sqrt{n}}\right) = \mathbb{P}\left(\sqrt{\mathrm{Var}(X_1)}G \geqslant y\right)$$

où $G \sim \mathcal{N}(0,1)$.

2. $G_i \sim \mathcal{N}(0,1)$ donc $\tilde{G}_n \sim \mathcal{N}(0,n)$.

$$p_n(x) = \mathbb{P}(\tilde{G}_n \geqslant nx) = \mathbb{P}(\sqrt{n}G_1 \geqslant nx) = \mathbb{P}(G_1 \geqslant \sqrt{n}x) = \int_{\sqrt{n}x}^{+\infty} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy$$

Or,

$$\ln \int_y^{+\infty} e^{-\frac{z^2}{2}} dz \sim -\frac{y^2}{2}$$

Car d'un côté on a,

$$\int_y^{+\infty} e^{-\frac{z^2}{2}} dz \leqslant \int_y^{+\infty} \frac{z}{y} e^{-\frac{z^2}{2}} dz = \frac{1}{2} e^{-\frac{y^2}{2}}$$

et de l'autre il suffit de minorer $\int_y^{+\infty} \ldots \geqslant \int_y^{2y} \ldots$. Finalement on a donc :

$$\ln(p_n(x)) \sim \frac{-nx^2}{2}$$

3.

4.

5. $Y \sim p$

$$\mathbb{E}(\mathbb{1}_{Z \geqslant x} \frac{p}{q}(z)) = \int_x^{+\infty} \frac{p}{q}(z) q(z) dz = \int_x^{+\infty} p(z) dz = \mathbb{P}(Y \geqslant x)$$

Le q optimal est (cf cours) :

$$q(z) = \frac{\mathbb{1}_{Z \geqslant x} p(z)}{\mathbb{P}(Y \geqslant x)}$$

6. Estimateur naïf :

$$\frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\frac{\tilde{X}_n^k}{n} \geqslant x} \xrightarrow[K \to +\infty]{} \mathbb{P}\left(\frac{\tilde{X}_n}{n} \geqslant x\right) = p_n(x)$$

et le TCL nous donne que :

$$\frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\frac{\tilde{X}_n^k}{n} \geqslant x} \in p_n(x) \pm \frac{\sqrt{\operatorname{Var}\left(\mathbb{1}_{\frac{\tilde{X}_n}{n} \geqslant x}\right)}}{\sqrt{n}}$$

On considère l'erreur relative,

$$\frac{\sqrt{\operatorname{Var}\left(\mathbb{1}_{\frac{\tilde{X}_n}{n} \geqslant x}\right)}}{p_n(x)} = \frac{\sqrt{p_n(x)(1 - p_n(x))}}{p_n(x)} \sim \frac{1}{\sqrt{p_n(x)}}$$

car $\frac{\tilde{X}_n}{n} \sim \mathcal{B}(p_n(x))$.

7. $L_n \colon \mathbb{R} \longrightarrow \mathbb{R}_+^*$ telle que $\mathbb{E}\left(\frac{1}{L_n}(\tilde{X}_n)\right) = 1$ et :

$$\mathbb{E}(\varphi(\tilde{X}_n)) = \mathbb{E}[\varphi(\tilde{Y}_n) L_n(\tilde{X}_n)] \quad \forall \varphi \text{ mesurable positive}$$

$$\varphi(y) = \frac{\mathbb{1}_B(y)}{L_n(y)} \qquad \mathbb{P}(\tilde{T}_n \in B) = \mathbb{E}\left(\frac{\mathbb{1}_B(\tilde{X}_n)}{L_n(\tilde{X}_n)}\right)$$

$\tilde{I}_k = \frac{1}{K} \sum_{k=1}^{K} \mathbb{1}_{\frac{\tilde{Y}_n^k}{n} \geqslant x} L_n(\tilde{Y}_n^k)$.

$$\tilde{I}_K \xrightarrow[K \to +\infty]{} \mathbb{P}\left(\frac{\tilde{X}_n}{n} \geqslant x\right) = p_n(x)$$