

Méthodes de Réduction de Variance : Echantillonnage d'Importance

Gersende FORT

LTCI, CNRS / TELECOM ParisTech

I. Introduction

Exemple 1

On veut calculer

$$\mathcal{I} \stackrel{\text{def}}{=} \mathbb{P}(X \geq 2) = \mathbb{E} [\mathbb{1}_{[2, +\infty)}(X)]$$

$$X \sim f(x) = \frac{1}{\pi(1+x^2)} \quad (\text{Cauchy})$$

- L'estimateur de Monte Carlo classique

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_{[2, +\infty)}(X_k) \quad X_k, \text{ i.i.d. de même loi que } X$$

a pour variance $0.127/n$.

Exemple 1

$$\mathcal{I} \stackrel{\text{def}}{=} \mathbb{P}(X \geq 2) = \mathbb{E} \left[\mathbb{1}_{[2, +\infty)}(X) \right] \quad X \sim f(x) = \frac{1}{\pi(1+x^2)} \text{ (Cauchy)}$$

- En remarquant que

$$\begin{aligned} \int_2^{+\infty} \frac{1}{\pi(1+x^2)} dx &= 2 \int_0^{1/2} \frac{1}{2\pi(1+y^2)} dy && y = 1/x \\ &= \mathbb{E} \left[\frac{1}{2\pi(1+U^2)} \right] && U \sim \mathcal{U}([0, 1/2]) \end{aligned}$$

on peut aussi approcher \mathcal{I} par l'estimateur de Monte Carlo

$$\frac{1}{n} \sum_{k=1}^n \frac{1}{2\pi(1+U_k^2)} \quad U_k, \text{ i.i.d. de même loi que } U$$

qui a pour variance $0.951 \cdot 10^{-4}/n$

Exemple 1

$$\begin{aligned} \mathcal{I} &\stackrel{\text{def}}{=} \mathbb{P}(X \geq 2) = \mathbb{E} [\mathbb{1}_{[2, +\infty)}(X)] & X &\sim f(x) = \frac{1}{\pi(1+x^2)} \text{ (Cauchy)} \\ &= \mathbb{E} \left[\frac{1}{2\pi(1+U^2)} \right] & U &\sim \mathcal{U}([0, 1/2]) \end{aligned}$$

- La seconde approche permet de réduire la variance d'un facteur 1 300

► **Par un changement de loi**, on s'est ramené à un autre problème de Monte Carlo de variance plus faible

Exemple 2

On veut calculer

 $(c > 0)$

$$\mathcal{I} \stackrel{\text{def}}{=} \mathbb{P}(Z > c)$$

$$Z \sim \mathcal{N}(0,1).$$

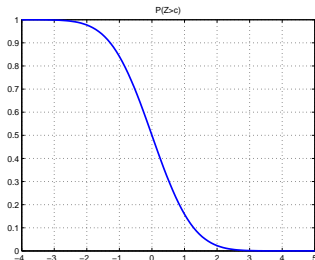
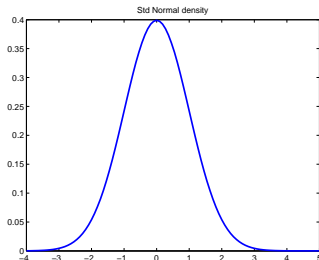


FIG.: [gauche] On trace la densité d'une loi $\mathcal{N}(0,1)$. [droite] On trace la probabilité $\mathbb{P}(Z > c)$ en fonction de c

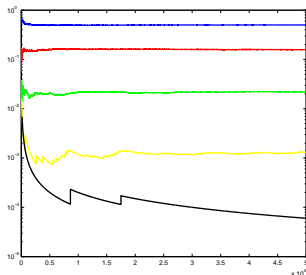
Lorsque $c = 8$, des calculateurs standard retournent $\mathcal{I} = 0$ - calculée sous Matlab, via la fonction "erf" - alors que $\mathbb{P}(Z > 8) > 0$

Approximation par méthode de Monte Carlo (classique)

- Calculer :

$$\bar{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{Z_k > c}$$

où Z_1, \dots, Z_n sont des
v.a. i.i.d. de loi $\mathcal{N}(0,1)$



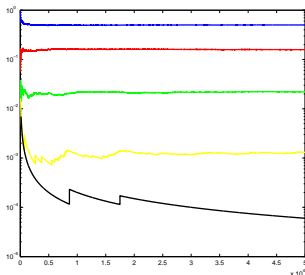
$\bar{\mu}_n$ pour $c \in \{0, 1, \dots, 4\}$

Approximation par méthode de Monte Carlo (classique)

- Calculer :

$$\bar{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n \mathbb{I}_{Z_k > c}$$

où Z_1, \dots, Z_n sont des
v.a. i.i.d. de loi $\mathcal{N}(0,1)$



$\bar{\mu}_n$ pour $c \in \{0, 1, \dots, 4\}$

- La variance de l'estimateur de Monte Carlo naïf est σ^2/n où $\sigma^2 = \mathcal{I}(1 - \mathcal{I})$.
- Erreur relative : (inégalité vraie car $c \geq 0$)

$$\frac{\sigma}{\bar{\mathcal{I}}} = \frac{\sqrt{\mathcal{I}(1 - \mathcal{I})}}{\bar{\mathcal{I}}} = \frac{\sqrt{1 - \bar{\mathcal{I}}}}{\sqrt{\bar{\mathcal{I}}}} \geq \frac{1}{\sqrt{2}} \frac{1}{\sqrt{\bar{\mathcal{I}}}} \rightarrow +\infty \quad \text{quand } \mathcal{I} \rightarrow 0$$

Changement de loi

Pour tout $\theta \in \mathbb{R}$, on a

$$\mathbb{P}(Z > c) = \mathbb{E} [\mathbf{1}_{Z+\theta > c} \exp(-0.5\theta^2 - \theta Z)] = \mathbb{E}_\theta [\mathbf{1}_{Z > c} \exp(0.5\theta^2 - \theta Z)]$$

↔ Quelle est la valeur *optimale* de θ ?

Changement de loi

Pour tout $\theta \in \mathbb{R}$, on a

$$\mathbb{P}(Z > c) = \mathbb{E} [\mathbb{1}_{Z+\theta > c} \exp(-0.5\theta^2 - \theta Z)] = \mathbb{E}_\theta [\mathbb{1}_{Z > c} \exp(0.5\theta^2 - \theta Z)]$$

↔ Quelle est la valeur *optimale* de θ ?

Si existence : prendre θ qui minimise la variance de l'estimateur de Monte Carlo
i.e.

$$\theta \text{ qui minimise } \sigma^2(\theta) \stackrel{\text{def}}{=} \mathbb{E}_\theta [\mathbb{1}_{Z > c} \exp(\theta^2 - 2\theta Z)] = \mathbb{E} [\mathbb{1}_{Z > c} \exp(0.5\theta^2 - \theta Z)]$$

Changement de loi

Pour tout $\theta \in \mathbb{R}$, on a

$$\mathbb{P}(Z > c) = \mathbb{E} [\mathbb{1}_{Z+\theta > c} \exp(-0.5\theta^2 - \theta Z)] = \mathbb{E}_\theta [\mathbb{1}_{Z > c} \exp(0.5\theta^2 - \theta Z)]$$

↔ Quelle est la valeur optimale de θ ?

Si existence : prendre θ qui minimise la variance de l'estimateur de Monte Carlo
i.e.

$$\theta \text{ qui minimise } \sigma^2(\theta) \stackrel{\text{def}}{=} \mathbb{E}_\theta [\mathbb{1}_{Z > c} \exp(\theta^2 - 2\theta Z)] = \mathbb{E} [\mathbb{1}_{Z > c} \exp(0.5\theta^2 - \theta Z)]$$

- $\theta \mapsto \sigma^2(\theta)$ est une fonction convexe de classe C^2 , et $\lim_{|\theta| \rightarrow \infty} \sigma^2(\theta) = \infty$. \Rightarrow
Unique minimum θ_*
- θ_* est l'unique solution de $d\sigma^2(\theta)/d\theta = 0$ i.e.

$$\mathbb{E} [\mathbb{1}_{Z > c} (\theta - Z) \exp(-\theta Z)] = 0.$$

Changement de loi optimal

$$\frac{d\sigma^2(\theta)}{d\theta} = \exp(0.5\theta^2) \mathbb{E}[\mathbb{1}_{Z>c} (\theta - Z) \exp(-\theta Z)].$$

- $\theta = 0$ n'est pas la solution optimale. Pourquoi?
- La solution optimale θ_* vérifie $\theta_* > c$. Pourquoi?

Changement de loi optimal

$$\frac{d\sigma^2(\theta)}{d\theta} = \exp(0.5\theta^2) \mathbb{E}[\mathbb{1}_{Z>c} (\theta - Z) \exp(-\theta Z)].$$

- $\theta = 0$ n'est pas la solution optimale. Pourquoi?
- La solution optimale θ_* vérifie $\theta_* > c$. Pourquoi?

Par des calculs standard, on a

$$\mathbb{E}[\mathbb{1}_{Z>c} (\theta - Z) \exp(-\theta Z)] = 2\theta \mathbb{P}(Z > c + \theta) - \frac{1}{\sqrt{2\pi}} \exp(-0.5(c + \theta)^2)$$

Mais pas de solution explicite qui, de plus, nécessiterait le calcul d'une quantité semblable à celle **inconnue!**.

Changement de loi optimal

$$\frac{d\sigma^2(\theta)}{d\theta} = \exp(0.5\theta^2) \mathbb{E}[\mathbb{1}_{Z>c} (\theta - Z) \exp(-\theta Z)].$$

- $\theta = 0$ n'est pas la solution optimale. Pourquoi?
- La solution optimale θ_* vérifie $\theta_* > c$. Pourquoi?

Par des calculs standard, on a

$$\mathbb{E}[\mathbb{1}_{Z>c} (\theta - Z) \exp(-\theta Z)] = 2\theta \mathbb{P}(Z > c + \theta) - \frac{1}{\sqrt{2\pi}} \exp(-0.5(c + \theta)^2)$$

Mais pas de solution explicite qui, de plus, nécessiterait le calcul d'une quantité semblable à celle **inconnue!**.

Lorsque c est grand,

$$2\theta \mathbb{P}(Z > c + \theta) - \frac{1}{\sqrt{2\pi}} \exp(-0.5(c + \theta)^2) \sim \frac{1}{\sqrt{2\pi}} \exp(-0.5(c + \theta)^2) \left\{ \frac{2\theta}{c + \theta} - 1 \right\}$$

⇒ choisir θ (très) proche de c .

Application numérique, lorsque $c = 8$

- l'estimateur de Monte Carlo usuel $\bar{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{Z_k > c}$
- l'estimateur d'échantillonnage d'importance

$$\hat{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{Z_k + \theta > c} \exp(-0.5\theta^2 - \theta Z_k)$$

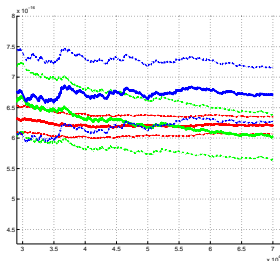
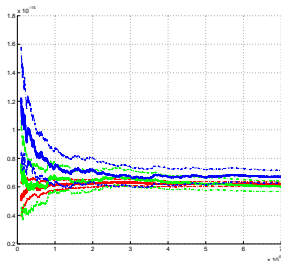


FIG.: $n \mapsto \hat{\mu}_n$ lorsque $\theta = c$ (rouge), $\theta = c + 2$ (en vert) et $\theta = c - 2$ (en bleu). En pointillés, les IC à 95%

Résultat : $\bar{\mu}_n = 0$ $\hat{\mu}_n = 6.25 \cdot 10^{-16}$ lorsque $n = 50\,000$.

II. Etude de l'estimateur d'échantillonnage d'importance

Relation fondamentale

- Pour simplifier l'exposé, on suppose que les lois des v.a. ont des densités par rapport à la mesure de Lebesgue sur \mathbb{R}^d .
- Sous \mathbb{P}_f , les v.a. sont i.i.d. de loi de densité f .

On souhaite calculer

$$\mathcal{I} \stackrel{\text{def}}{=} \mathbb{E}_f [\phi(Z)].$$

La méthode d'échantillonnage d'importance repose sur la relation fondamentale

$$\mathbb{E}_f [\phi(Z)] = \mathbb{E}_g \left[\phi(Z) \frac{f(Z)}{g(Z)} \right]$$

pour toute densité g telle que $\text{Supp}(f\phi) \subseteq \text{Supp}(g)$. ↔ (preuve)

Estimateurs

On compare l'estimateur d'échantillonnage d'importance

$$\hat{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n \frac{f(Z_k)}{g(Z_k)} \phi(Z_k)$$

où les v.a. $\{Z_k, k \geq 0\}$ sont i.i.d. de loi g

à l'estimateur de Monte Carlo

$$\bar{\mu}_n \stackrel{\text{def}}{=} \frac{1}{n} \sum_{k=1}^n \phi(Y_k)$$

où les v.a. $\{Y_k, k \geq 0\}$ sont i.i.d. de loi f .

Hypothèses

On suppose que

- 1 $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ est mesurable et $\mathbb{E}_f[\phi^2(Z)] < +\infty$.
- 2 la densité instrumentale g est telle que $\text{Supp}(f\phi) \subseteq \text{Supp}(g)$.
- 3 $0 < \mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] - (\mathbb{E}_f[\phi(Z)])^2 < +\infty$.

On étudie

- le biais
- la variance
- le choix optimal de la loi instrumentale g .

Biais et Consistance

- L'estimateur d'échantillonnage d'importance est sans biais.
- Il est fortement consistant.

↪ (preuve)

Variance

La variance de l'estimateur d'échantillonnage d'importance est donnée par

$$\begin{aligned}\text{Var}_g(\hat{\mu}_n) &= \frac{1}{n} \mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] - \frac{1}{n} (\mathbb{E}_f[\phi(Z)])^2 \\ &= \text{Var}_f(\bar{\mu}_n) + \frac{1}{n} \mathbb{E}_f \left[\phi^2(Z) \left\{ \frac{f(Z)}{g(Z)} - 1 \right\} \right]\end{aligned}$$

↪ (preuve)

Variance

La variance de l'estimateur d'échantillonnage d'importance est donnée par

$$\begin{aligned}\text{Var}_g(\hat{\mu}_n) &= \frac{1}{n} \mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] - \frac{1}{n} (\mathbb{E}_f[\phi(Z)])^2 \\ &= \text{Var}_f(\bar{\mu}_n) + \frac{1}{n} \mathbb{E}_f \left[\phi^2(Z) \left\{ \frac{f(Z)}{g(Z)} - 1 \right\} \right]\end{aligned}$$

↪ (preuve)

- ▶ La variance est finie dès que $\mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] < +\infty$.
 - Une condition suffisante pour cela lorsque $\mathbb{E}_f[\phi^2(Z)] < +\infty$ est que $\sup_{\mathbb{R}^d} f/g < +\infty$. Si ce n'est pas le cas, les poids $f(Z_k)/g(Z_k)$ sont très variables ce qui fait que dans l'expression de $\hat{\mu}_n$, seuls quelques points sont significatifs → instabilité de l'estimateur. ▶ Exemples
 - En pratique, on choisit une loi instrumentale g qui a des queues plus lourdes que celles de f .

Variance

La variance de l'estimateur d'échantillonnage d'importance est donnée par

$$\begin{aligned}\text{Var}_g(\hat{\mu}_n) &= \frac{1}{n} \mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] - \frac{1}{n} (\mathbb{E}_f[\phi(Z)])^2 \\ &= \text{Var}_f(\bar{\mu}_n) + \frac{1}{n} \mathbb{E}_f \left[\phi^2(Z) \left\{ \frac{f(Z)}{g(Z)} - 1 \right\} \right]\end{aligned}$$

↔ (preuve)

► Cet estimateur est de variance plus faible que l'estimateur de Monte Carlo ssi

$$\mathbb{E}_f \left[\phi^2(Z) \left\{ \frac{f(Z)}{g(Z)} - 1 \right\} \right] < 0.$$

↔ le choix optimal du changement de loi g dépend de f et de ϕ .

Estimateur optimal

Par Jensen

$$\text{Var}_g (\hat{\mu}_n) \geq (\mathbb{E}_f [|\phi(Z)|])^{1/2},$$

avec égalité en

$$g_\star(z) \stackrel{\text{def}}{=} \frac{|\phi(z)|f(z)}{\mathbb{E}_f [|\phi(Z)|]}$$

auquel cas

$$\text{Var}_{g_\star} (\hat{\mu}_n) = 0 \quad !!!$$

↪ (preuve)

Estimateur optimal

Par Jensen

$$\text{Var}_g(\hat{\mu}_n) \geq (\mathbb{E}_f[|\phi(Z)|])^{1/2},$$

avec égalité en

$$g_\star(z) \stackrel{\text{def}}{=} \frac{|\phi(z)|f(z)}{\mathbb{E}_f[|\phi(Z)|]}$$

auquel cas

$$\text{Var}_{g_\star}(\hat{\mu}_n) = 0 \quad !!!$$

↪ (preuve)

- Ce changement de loi optimal ne peut pas être mis en oeuvre en pratique puisqu'il dépend de la quantité inconnue \mathcal{I} (si $\phi > 0$).
- En revanche, il exprime le fait que la variance est d'autant plus faible que la loi instrumentale g est choisie "proche de $f|\phi|$ ".
- En pratique, on choisit g dans une famille \mathcal{P} de densités; et on choisit (ou essaye de s'approcher de ...) la densité qui vérifie

$$g_\star = \operatorname{argmin}_{g \in \mathcal{P}} \text{Var}_g(\hat{\mu}_n).$$

Intervalles de confiance

Par le TCL pour des v.a. i.i.d.

$$\sqrt{n} (\hat{\mu}_n - \mathbb{E}_f [\phi(Z)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2)$$

où

$$\sigma^2 = \mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] - (\mathbb{E}_f [\phi(Z)])^2.$$

↔ (preuve)

Intervalles de confiance

Par le TCL pour des v.a. i.i.d.

$$\sqrt{n} (\hat{\mu}_n - \mathbb{E}_f [\phi(Z)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2)$$

où

$$\sigma^2 = \mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] - (\mathbb{E}_f [\phi(Z)])^2.$$

↔ (preuve)

Un intervalle de confiance à 95% pour l'estimation de $\mathbb{E}_f[\phi(Z)]$ est donné par

$$\left[\hat{\mu}_n - 1.96 \frac{\sigma}{\sqrt{n}}; \quad \hat{\mu}_n + 1.96 \frac{\sigma}{\sqrt{n}} \right]$$

Intervalles de confiance

Par le TCL pour des v.a. i.i.d.

$$\sqrt{n} (\hat{\mu}_n - \mathbb{E}_f [\phi(Z)]) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2)$$

où

$$\sigma^2 = \mathbb{E}_f \left[\phi^2(Z) \frac{f(Z)}{g(Z)} \right] - (\mathbb{E}_f [\phi(Z)])^2.$$

↪ (preuve)

Un intervalle de confiance à 95% pour l'estimation de $\mathbb{E}_f[\phi(Z)]$ est donné par

$$\left[\hat{\mu}_n - 1.96 \frac{\sigma}{\sqrt{n}}; \quad \hat{\mu}_n + 1.96 \frac{\sigma}{\sqrt{n}} \right]$$

Lorsque σ n'est pas connu : on l'estime par un estimateur fortement consistant et le théorème de Slutsky permet de montrer que l'IC ci-dessus reste valable en remplaçant σ par son estimation.

Estimateur auto-normalisé

L'estimateur d'échantillonnage d'importance peut être substitué par sa version auto-normalisée

$$\frac{1}{\sum_{j=1}^n \frac{f(Z_j)}{g(Z_j)}} \sum_{k=1}^n \frac{f(Z_k)}{g(Z_k)} \phi(Z_k)$$

- Cet estimateur est plus robuste aux grandes valeurs du ratio f/g .
- Il ne nécessite la connaissance de f qu'à une constante multiplicative près.
- Il est fortement consistant.
- Par application de la delta-méthode, sa variance est donnée par

$$\mathbb{E}_f \left[\{ \phi(Z) - \mathbb{E}_f [\phi(Z)] \}^2 \frac{f(Z)}{g(Z)} \right]$$

III. Exemples de changements de lois

Introduction

La densité instrumentale est choisie dans une famille de densités \mathcal{P} , comme la densité qui optimise un critère **en général lié à la variance** de l'estimateur.

Nous considérerons trois exemples :

- ④ dans le premier exemple, g est choisie comme la densité qui minimise un majorant de la variance.

Introduction

La densité instrumentale est choisie dans une famille de densités \mathcal{P} , comme la densité qui optimise un critère **en général lié à la variance** de l'estimateur.

Nous considèrerons trois exemples :

- 1 dans le premier exemple, g est choisie comme la densité qui minimise un majorant de la variance.
- 2 dans le second exemple, f est la densité $\mathcal{N}_d(0, \text{Id})$ et

$$\mathcal{P} = \{\mathcal{N}_d(\theta, \text{Id}), \quad \theta \in \mathbb{R}^d\}.$$

Dans ce cas (+ hyp sur la fonction ϕ), on peut montrer que la variance de l'estimateur est une fonction convexe qui possède un unique minimum.

Introduction

La densité instrumentale est choisie dans une famille de densités \mathcal{P} , comme la densité qui optimise un critère **en général lié à la variance** de l'estimateur.

Nous considérerons trois exemples :

- ① dans le premier exemple, g est choisie comme la densité qui minimise un majorant de la variance.
- ② dans le second exemple, f est la densité $\mathcal{N}_d(0, \text{Id})$ et

$$\mathcal{P} = \{\mathcal{N}_d(\theta, \text{Id}), \quad \theta \in \mathbb{R}^d\}.$$

Dans ce cas (+ hyp sur la fonction ϕ), on peut montrer que la variance de l'estimateur est une fonction convexe qui possède un unique minimum.

- ③ dans le dernier exemple, $d = 1$ et la famille \mathcal{P} est

$$\mathcal{P} = \{\exp(\theta x - \psi(\theta)) f(x), \quad \theta \in \mathbb{R}\} \quad \psi(\theta) \stackrel{\text{def}}{=} \log \int \exp(\theta x) f(x) dx.$$

(lorsque f est gaussienne, le cas 3 coïncide avec le cas 2)

Exemple 1

On souhaite calculer

$$\mathcal{I} \stackrel{\text{def}}{=} \int_a^{+\infty} z^{\alpha-1} \exp(-z) dz, \quad \alpha > 1, a > 0.$$

On re-écrit \mathcal{I} sous la forme $\mathbb{E}_f[\phi(Z)]$ avec

$$\phi(z) = z^{\alpha-1} \mathbb{1}_{[a, +\infty)}(z) \quad f(z) = \exp(-z) \mathbb{1}_{\mathbb{R}^+}(z).$$

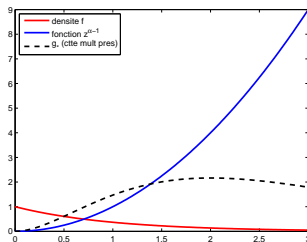


FIG.: Fonction $z \mapsto z^{\alpha-1}$; Densité f ; et $g_*(z) \propto f(z)z^{\alpha-1}$

($\alpha = 3$)

Changement de loi possible

On cherche le changement de lois parmi la famille des lois exponentielles translatées

$$\mathcal{P} \stackrel{\text{def}}{=} \{g_{\theta}(z) = \theta \exp(-\theta(z - a)) \mathbb{I}_{[a, +\infty)}(z), \quad \theta > 0\}$$

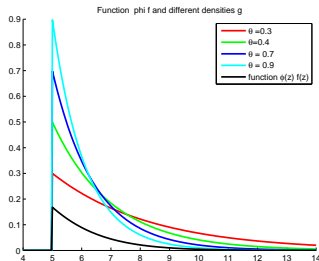


FIG.: Densités $g_{\theta}(z) = \theta \exp(-\theta(z - a)) \mathbb{I}_{[a, +\infty)}(z)$ pour différentes valeurs de θ ($a = 5$)

Premier majorant

On choisit g comme minimisant le majorant de la variance

$$\int_{\text{Supp}(f\phi)} \phi^2(z) \frac{f^2(z)}{g_\theta(z)} dz \leq \left(\sup_{z \in \text{Supp}(f\phi)} \frac{f(z)}{g_\theta(z)} \right) \int_{\text{Supp}(f\phi)} \phi^2(z) f(z) dz$$

Premier majorant

On choisit g comme minimisant le majorant de la variance

$$\int_{\text{Supp}(f\phi)} \phi^2(z) \frac{f^2(z)}{g_\theta(z)} dz \leq \left(\sup_{z \in \text{Supp}(f\phi)} \frac{f(z)}{g_\theta(z)} \right) \int_{\text{Supp}(f\phi)} \phi^2(z) f(z) dz$$

- Dans cet exemple $\text{Supp}(f\phi) = [a, +\infty[$ et

$$\begin{aligned} \sup_{z \geq a} \frac{f(z)}{g_\theta(z)} &= \theta^{-1} \sup_{z \geq a} \exp(z(\theta - 1) - a\theta) \\ &= \begin{cases} \theta^{-1} \exp(-a) & \text{si } \theta \in (0,1] \\ +\infty & \text{si } \theta > 1 \end{cases} \end{aligned}$$

Rmq : $\theta \leq 1 \Rightarrow$ la densité instrumentale g_θ a des queues plus lourdes (ou aussi lourdes) que f .

Premier majorant

On choisit g comme minimisant le majorant de la variance

$$\int_{\text{Supp}(f\phi)} \phi^2(z) \frac{f^2(z)}{g_\theta(z)} dz \leq \left(\sup_{z \in \text{Supp}(f\phi)} \frac{f(z)}{g_\theta(z)} \right) \int_{\text{Supp}(f\phi)} \phi^2(z) f(z) dz$$

- Dans cet exemple $\text{Supp}(f\phi) = [a, +\infty[$ et

$$\begin{aligned} \sup_{z \geq a} \frac{f(z)}{g_\theta(z)} &= \theta^{-1} \sup_{z \geq a} \exp(z(\theta - 1) - a\theta) \\ &= \begin{cases} \theta^{-1} \exp(-a) & \text{si } \theta \in (0, 1] \\ +\infty & \text{si } \theta > 1 \end{cases} \end{aligned}$$

Rmq : $\theta \leq 1 \Rightarrow$ la densité instrumentale g_θ a des queues plus lourdes (ou aussi lourdes) que f .

- On minimise ce majorant en θ . On obtient $\theta_\star^{(1)} = 1$ i.e. $g_\star^{(1)}$ est la loi f translatée.

Premier majorant

On choisit g comme minimisant le majorant de la variance

$$\int_{\text{Supp}(f\phi)} \phi^2(z) \frac{f^2(z)}{g_\theta(z)} dz \leq \left(\sup_{z \in \text{Supp}(f\phi)} \frac{f(z)}{g_\theta(z)} \right) \int_{\text{Supp}(f\phi)} \phi^2(z) f(z) dz$$

- Dans cet exemple $\text{Supp}(f\phi) = [a, +\infty[$ et

$$\begin{aligned} \sup_{z \geq a} \frac{f(z)}{g_\theta(z)} &= \theta^{-1} \sup_{z \geq a} \exp(z(\theta - 1) - a\theta) \\ &= \begin{cases} \theta^{-1} \exp(-a) & \text{si } \theta \in (0, 1] \\ +\infty & \text{si } \theta > 1 \end{cases} \end{aligned}$$

Rmq : $\theta \leq 1 \Rightarrow$ la densité instrumentale g_θ a des queues plus lourdes (ou aussi lourdes) que f .

- On minimise ce majorant en θ . On obtient $\theta_\star^{(1)} = 1$ i.e. $g_\star^{(1)}$ est la loi f translatée.
- Le changement de loi g_θ est choisi d'après un critère
 - qui assure f/g_θ borné.
 - qui dépend de f mais pas de ϕ .

Second majorant

On choisit g comme minimisant le majorant de la variance

$$\int_{\text{Supp}(f\phi)} \phi^2(z) \frac{f^2(z)}{g\theta(z)} dz \leq \left(\sup_{z \in \text{Supp}(f\phi)} \phi(z) \frac{f(z)}{g\theta(z)} \right) \int_{\text{Supp}(f\phi)} \phi(z) f(z) dz$$

Second majorant

On choisit g comme minimisant le majorant de la variance

$$\int_{\text{Supp}(f\phi)} \phi^2(z) \frac{f^2(z)}{g\theta(z)} dz \leq \left(\sup_{z \in \text{Supp}(f\phi)} \phi(z) \frac{f(z)}{g\theta(z)} \right) \int_{\text{Supp}(f\phi)} \phi(z) f(z) dz$$

- Dans cet exemple

$$\sup_{z \geq a} \phi(z) \frac{f(z)}{g\theta(z)} = \theta^{-1} \sup_{z \geq a} \exp(\theta(z-a) - z) z^{\alpha-1}$$

et la majorant n'est fini que si $\theta \in]0, 1[$. On obtient

$$\theta_{\star}^{(2)} = \frac{(a - \alpha) + \sqrt{(a - \alpha)^2 + 4a}}{2a}$$

Lorsque $a = 5$ et $\alpha = 3$, on a $\theta_{\star}^{(2)} = 0.69$

Second majorant

On choisit g comme minimisant le majorant de la variance

$$\int_{\text{Supp}(f\phi)} \phi^2(z) \frac{f^2(z)}{g_\theta(z)} dz \leq \left(\sup_{z \in \text{Supp}(f\phi)} \phi(z) \frac{f(z)}{g_\theta(z)} \right) \int_{\text{Supp}(f\phi)} \phi(z) f(z) dz$$

- Dans cet exemple

$$\sup_{z \geq a} \phi(z) \frac{f(z)}{g_\theta(z)} = \theta^{-1} \sup_{z \geq a} \exp(\theta(z-a) - z) z^{\alpha-1}$$

et la majorant n'est fini que si $\theta \in]0, 1[$. On obtient

$$\theta_\star^{(2)} = \frac{(a - \alpha) + \sqrt{(a - \alpha)^2 + 4a}}{2a}$$

Lorsque $a = 5$ et $\alpha = 3$, on a $\theta_\star^{(2)} = 0.69$

- Le changement de loi g_θ est choisi d'après un critère qui dépend de f et de ϕ .

Comparaison numérique

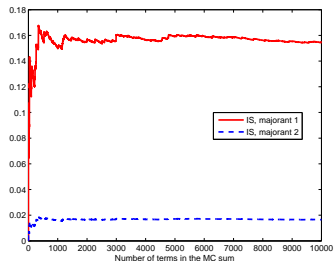
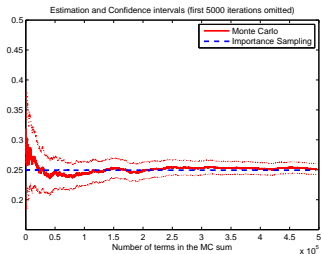


FIG.: [gauche] On trace l'évolution de l'estimateur de MC en fonction du nombre de termes n dans la somme; et l'évolution de l'estimateur d'importance dans le cas $\theta = \theta_\star^{(2)}$. On trace aussi l'IC à 95%. [droite] On trace l'évolution de $\sqrt{n \text{Var}_{g_\theta}(\hat{\mu}_n)}$ pour $\theta \in \{\theta_\star^{(1)}, \theta_\star^{(2)}\}$

La méthode 2 permet de réduire la variance de l'estimateur MC d'un facteur 37 400; et améliore la réduction de variance obtenue par la méthode 1 d'un facteur 64.

Exemple 2 (cas Gaussien)

On souhaite calculer

$$\mathcal{I} \stackrel{\text{def}}{=} \mathbb{E}[\phi(Z)] \quad \text{où } Z \sim \mathcal{N}_d(0, \text{Id})$$

C'est la situation par exemple du call

- d'une option européenne ($d = 1$)

$$\exp(-rT) \mathbb{E} \left[\left(S_0 \exp([r - 0.5\sigma^2]T + \sigma\sqrt{T}Z) - K \right)_+ \right]$$

Exemple 2 (cas Gaussien)

On souhaite calculer

$$\mathcal{I} \stackrel{\text{def}}{=} \mathbb{E} [\phi(Z)] \quad \text{où } Z \sim \mathcal{N}_d(0, \text{Id})$$

C'est la situation par exemple du call

- d'une option européenne ($d = 1$)

$$\exp(-rT) \mathbb{E} \left[\left(S_0 \exp([r - 0.5\sigma^2]T + \sigma\sqrt{T}Z) - K \right)_+ \right]$$

- d'une option asiatique ($d \geq 1$)

$$\exp(-rT) \mathbb{E} \left[\left(\frac{1}{d} \sum_{k=1}^d S_0 \exp([r - 0.5\sigma^2] \frac{kT}{d} + \sigma\sqrt{\frac{T}{d}} \sum_{j=1}^k Z_j) - K \right)_+ \right]$$

avec $Z = (Z_1, \dots, Z_d)$.

↪ (preuve)

Changement de loi possible

On cherche un changement de loi dans la famille

$$\mathcal{P} \stackrel{\text{def}}{=} \{\mathcal{N}_d(\theta, \text{Id}), \quad \theta \in \mathbb{R}^d\}$$

On montre (cf. TD)

- que pour tout $\theta \in \mathbb{R}^d$

$$\mathcal{I} = \mathbb{E} [\phi(Z + \theta) \exp(-0.5\theta'\theta - \theta'Z)]$$

- la variance de l'estimateur MC du terme de droite est $\sigma^2(\theta)/n$

$$\sigma^2(\theta) \stackrel{\text{def}}{=} \text{Var} [\phi(Z + \theta) \exp(-0.5\theta'\theta - \theta'Z)];$$

$\theta \mapsto \sigma^2(\theta)$ est une fonction convexe qui possède un unique minimum solution de

$$\mathbb{E} [(\theta - Z) \exp(-\theta'Z) \phi^2(Z)] = 0.$$

Exemple, cas $d = 1$

$$\mathcal{I} \stackrel{\text{def}}{=} \exp(-rT) \mathbb{E} [(\lambda \exp(\sigma Z) - K)_+] \quad (\text{calculable explicitement par Black-Scholes})$$

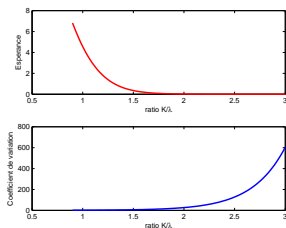


FIG.: On trace la valeur de \mathcal{I} et du coefficient de variation en fonction du ratio K/λ . λ est fixé et on fait varier le strike K dans l'intervalle $[0.9\lambda, 3\lambda]$

$$T = 1 \quad S_0 = 40 \quad \sigma = 0.25 \quad r = 0.05$$

$$\lambda = S_0 \exp((r - 0.5\sigma^2)T)$$

Lorsque $K/\lambda = 3.68$, $\mathcal{I} = 6.26e - 07$.
Calculé avec $n = 10^6$ tirages :

- $\bar{\mu}_n = 0$.
- $\hat{\mu}_n = 6.26e - 07$,
 $\theta = \ln(K/\lambda)/(\sigma\sqrt{T})$
 $IC_{0.95} = [6.21e - 07; 6.33e - 07]$.

Exemple, cas $d = 1$

$$\mathcal{I} \stackrel{\text{def}}{=} \exp(-rT) \mathbb{E} [(\lambda \exp(\sigma Z) - K)_+] \quad (\text{calculable explicitement par Black-Scholes})$$

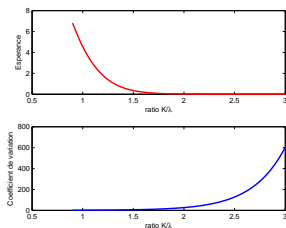


FIG.: On trace la valeur de \mathcal{I} et du coefficient de variation en fonction du ratio K/λ . λ est fixé et on fait varier le strike K dans l'intervalle $[0.9\lambda, 3\lambda]$

$$T = 1 \quad S_0 = 40 \quad \sigma = 0.25 \quad r = 0.05$$

$$\lambda = S_0 \exp((r - 0.5\sigma^2)T)$$

Lorsque $K/\lambda = 3.68$, $\mathcal{I} = 6.26e - 07$.
Calculé avec $n = 10^6$ tirages :

- $\bar{\mu}_n = 0$.
- $\hat{\mu}_n = 6.26e - 07$,
 $\theta = \ln(K/\lambda)/(\sigma\sqrt{T})$
 $IC_{0.95} = [6.21e - 07; 6.33e - 07]$.

On déduit des résultats précédents que $\theta \mapsto \sigma^2(\theta)$ est décroissante sur $]-\infty; \sigma^{-1} \log(K/\lambda)]$.

Choix de θ basé sur la technique de *linéarisation*

$$(\text{Rappel: } \mathbb{E}[\phi(Z)] = \mathbb{E}[\phi(Z + \theta) \exp(-0.5\theta'\theta - \theta'Z)])$$

Lorsque $\phi \geq 0$, on peut toujours écrire

$$\mathbb{E}[\phi(Z)] = \mathbb{E}[\exp(F(Z))] = \mathbb{E}[\exp(F(Z + \theta)) \exp(-0.5\theta'\theta - \theta'Z)]$$

avec

$$F(z) = \begin{cases} \ln \phi(z) & \text{si } \phi(z) > 0 \\ -\infty & \text{si } \phi(z) = 0 \end{cases}$$

Choix de θ basé sur la technique de *linéarisation*

$$(\text{Rappel: } \mathbb{E}[\phi(Z)] = \mathbb{E}[\phi(Z + \theta) \exp(-0.5\theta'\theta - \theta'Z)])$$

Lorsque $\phi \geq 0$, on peut toujours écrire

$$\mathbb{E}[\phi(Z)] = \mathbb{E}[\exp(F(Z))] = \mathbb{E}[\exp(F(Z + \theta)) \exp(-0.5\theta'\theta - \theta'Z)]$$

avec

$$F(z) = \begin{cases} \ln \phi(z) & \text{si } \phi(z) > 0 \\ -\infty & \text{si } \phi(z) = 0 \end{cases}$$

En écrivant

$$\exp(F(z + \theta)) \approx \exp(F(\theta) + z'\nabla F(\theta))$$

on a

$$\mathbb{E}[\phi(Z)] \approx \exp(F(\theta) - 0.5\theta'\theta) \mathbb{E}[\exp(Z'\{\nabla F(\theta) - \theta\})]$$

$$\implies \text{Prendre } \theta \text{ solution de } \theta = \nabla F(\theta)$$

\hookrightarrow Réduction de variance d'autant plus importante que l'approximation linéaire est correcte.

Exemple 3

Quelle que soit la densité f sur \mathbb{R} , on cherche la densité g dans la famille

$$\mathcal{P} \stackrel{\text{def}}{=} \{g_\theta(z) \propto \exp(\theta z)f(z), \quad \theta \in \mathbb{R}\}.$$

- La constante de normalisation de g_θ est liée à la **fonction génératrice des cumulants** ψ définie par

$$\psi(\theta) = \ln \int \exp(\theta z)f(z)dz.$$

- Lorsque $\theta > 0$, ces **"tilted"** densités déplacent la moyenne de la loi vers la droite.
- Dans le cas où $f = \mathcal{N}(0,1)$, cette méthode coïncide avec la précédente puisque $g_\theta = \mathcal{N}(\theta,1)$

↔ (cf. TP pour un exemple d'application)

Conclusion

Quel changement de loi pour approcher

$$\mathbb{E}[\phi(Z)]$$

Chercher un changement de loi parmi une famille de lois $\mathcal{P} = \{g_\theta, \theta \in \Theta\}$

- 1 critère de sélection de la loi dans cette famille, indépendant de ϕ ou pas.

Conclusion

Quel changement de loi pour approcher

$$\mathbb{E}[\phi(Z)]$$

Chercher un changement de loi parmi une famille de lois $\mathcal{P} = \{g_\theta, \theta \in \Theta\}$

- 1 critère de sélection de la loi dans cette famille, indépendant de ϕ ou pas.
- 2 Nouvelle génération : **Méthodes adaptatives**
 - Critère de sélection basé sur l'optimisation d'un critère non calculable en pratique, mais qui peut se re-écrire sous la forme d'une espérance.

Conclusion

Quel changement de loi pour approcher

$$\mathbb{E}[\phi(Z)]$$

Chercher un changement de loi parmi une famille de lois $\mathcal{P} = \{g_\theta, \theta \in \Theta\}$

- 1 critère de sélection de la loi dans cette famille, indépendant de ϕ ou pas.
- 2 Nouvelle génération : **Méthodes adaptatives**
 - Critère de sélection basé sur l'optimisation d'un critère non calculable en pratique, mais qui peut se re-écrire sous la forme d'une espérance.
Ex. prendre $g \in \mathcal{P}$ qui minimise la distance de Kullback

$$g_\star = \operatorname{argmin}_{g \in \mathcal{P}} \int \log \frac{f(z)}{g(z)} f(z) dz.$$

- On approche cette espérance par une somme de Monte Carlo type échantillonnage d'importance avec le changement de loi $g_{\theta^{(t)}}$.
- On optimise le critère approché pour déterminer $\theta^{(t+1)}$.
- On répète le mécanisme jusqu'à convergence.

Conclusion

Quel changement de loi pour approcher

$$\mathbb{E}[\phi(Z)]$$

Chercher un changement de loi parmi une famille de lois $\mathcal{P} = \{g_\theta, \theta \in \Theta\}$

- 1 critère de sélection de la loi dans cette famille, indépendant de ϕ ou pas.
- 2 Nouvelle génération : **Méthodes adaptatives**

Références

- Lemaire, V. and Pagès, G. *Unconstrained recursive importance sampling*, Technical report, 2008.
- Capriotti, L. *Least Squares Importance Sampling for Monte Carlo Security Pricing*, Quantitative Finance, 8(5), 2008, pages 485 - 497.
- Jourdain, B. et Lelong, J. *Robust Adaptive Importance Sampling for Normal Random Vectors*, Annals of Applied Probability, 19(5), 2009, 1687-1718
- Cappé, O., Douc, R., Gullin, A., Marin, J.-M. and Robert, C.P. *Adaptive Importance Sampling in General Mixture Classes*, Statistics and Computing 18, 2009, 447-459.

IV. Application au calcul d'options

Application au calcul du prix d'options européennes

On veut calculer

$$\mathcal{I} = \exp(-rT) \mathbb{E} \left[\left(S_0 \exp\{(r - 0.5\sigma^2)T + \sigma\sqrt{T}Z\} - K \right)_+ \right] \quad Z \sim \mathcal{N}(0,1)$$

avec

$$S_0 = 100 \quad K = 120 \quad r = 0.02 \quad \sigma = 0.2 \quad T = 10.$$

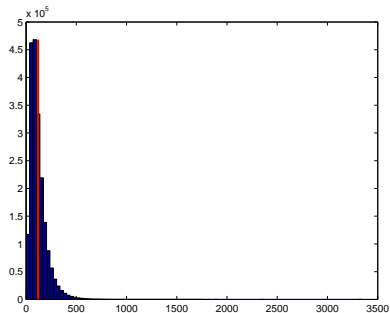
Nous allons comparer l'estimateur de MC usuel, à l'estimateur d'échantillonnage d'importance dans lequel θ est fixé à la valeur optimale.

Dans cet exemple, \mathcal{I} est de la forme

$$\mathcal{I} = \exp(-rT) \mathbb{E} [(\psi(Z) - K)_+]$$

et se calcule explicitement: $\mathcal{I} = 25.48$.

- On tire $N = 2e06$ v.a. $Z_k \sim \mathcal{N}(0,1)$ et on trace l'historgramme des valeurs $\psi(Z_k)$.
- On trace la valeur du strike K .



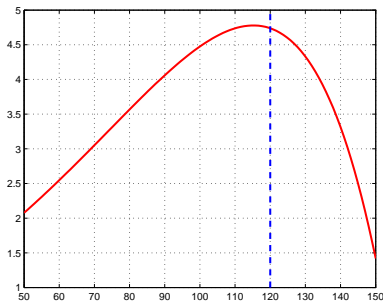
La probabilité que $Z \sim \mathcal{N}(0,1)$ soit tel que $\psi(Z) - K > 0$ est d'autant plus faible que K est grand (S_0, r, T, σ fixés). En effet

$$\mathbb{P}(\psi(Z) - K > 0) = \mathbb{P}\left(Z > \frac{\ln(K/S_0) - (r - 0.5\sigma^2)T}{\sigma\sqrt{T}}\right) = p(K)$$

- Pour obtenir 1 tirage tel que $\psi(Z) - K > 0$, il faut attendre $1/p(K)$ tirages en moyenne.
- La méthode d'échantillonnage d'importance est donc d'autant plus intéressante (quand comparée à MC standard) que K/S_0 est grand.

On trace l'évolution de la variance normalisée par l'espérance au carré, en fonction de K :

$$\frac{\mathbb{E} [(\psi(Z) - K)_+^2]}{\{\mathbb{E} [(\psi(Z) - K)_+]\}^2}$$



Dans la suite, $K = 120$.

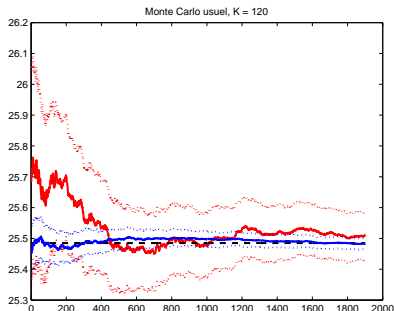
Estimateur de Monte Carlo usuel

Objectif : estimer $\mathcal{I} = \mathbb{E}[\phi(Z)]$

- Tirer N v.a. $Z_k \sim \mathcal{N}(0,1)$
- On trace (en rouge) l'évolution

$$n \mapsto \frac{1}{n} \sum_{k=1}^n \phi(Z_k)$$

et celle de l'intervalle de confiance à 95%. (échelle 10^{-3} sur les graphes en abscisse)



Estimateur d'échantillonnage d'importance

- Pour tout $\theta \in \mathbb{R}$, on a

$$\mathbb{E}[\phi(Z)] = \exp(-0.5\theta^2) \mathbb{E}[\phi(Z + \theta) \exp(-\theta Z)]$$

- La variance de l'estimateur d'échantillonnage d'importance lorsque la loi instrumentale est $g_\theta = \mathcal{N}(\theta, 1)$ est fonction de θ .

Comment choisir θ ?

- (Méthode 1) comme la solution de

$$\nabla[\ln \phi](\theta) = \theta \quad \theta \in \{z, \phi(z) > 0\}.$$

Estimateur d'échantillonnage d'importance

- Pour tout $\theta \in \mathbb{R}$, on a

$$\mathbb{E}[\phi(Z)] = \exp(-0.5\theta^2) \mathbb{E}[\phi(Z + \theta) \exp(-\theta Z)]$$

- La variance de l'estimateur d'échantillonnage d'importance lorsque la loi instrumentale est $g_\theta = \mathcal{N}(\theta, 1)$ est fonction de θ .

Comment choisir θ ?

- (Méthode 1) comme la solution de

$$\nabla[\ln \phi](\theta) = \theta \quad \theta \in \{z, \phi(z) > 0\}.$$

- (Méthode 2) comme la valeur qui minimise la variance (min unique) calculée par un algorithme de gradient stochastique.

Estimateur d'échantillonnage d'importance

- Pour tout $\theta \in \mathbb{R}$, on a

$$\mathbb{E}[\phi(Z)] = \exp(-0.5\theta^2) \mathbb{E}[\phi(Z + \theta) \exp(-\theta Z)]$$

- La variance de l'estimateur d'échantillonnage d'importance lorsque la loi instrumentale est $g_\theta = \mathcal{N}(\theta, 1)$ est fonction de θ .

Comment choisir θ ?

- (Méthode 1) comme la solution de

$$\nabla[\ln \phi](\theta) = \theta \quad \theta \in \{z, \phi(z) > 0\}.$$

- (Méthode 2) comme la valeur qui minimise la variance (min unique) calculée par un algorithme de gradient stochastique.
- (Méthode 3) ...

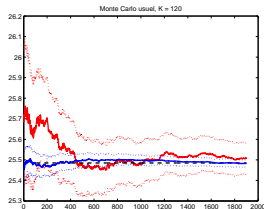
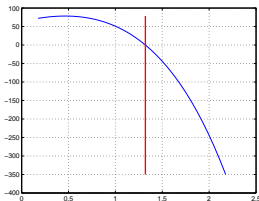
Choix de θ par Méthode 1

On calcule le drift par la recherche de la solution de $\nabla[\ln \phi](\theta) = \theta$ pour $\theta \in \{z, \phi(z) > 0\}$.

- (gauche) On trace la courbe $\theta \mapsto \nabla[\ln \phi](\theta) - \theta$ sur le domaine $\{z, \phi(z) > 0\}$. Elle possède un zéro unique: on trouve $\theta_{\star}^{(1)} = 1.319$.
- (droite) On trace (en bleu) l'évolution de

$$n \mapsto \frac{1}{n} \sum_{k=1}^n \phi(Z_k + \theta_{\star}^{(1)}) \exp(-\theta_{\star}^{(1)} Z_k) \exp(-0.5[\theta_{\star}^{(1)}]^2)$$

et de l'intervalle de confiance à 95%.

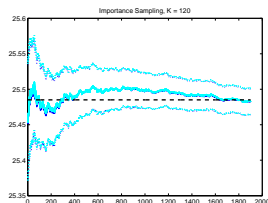
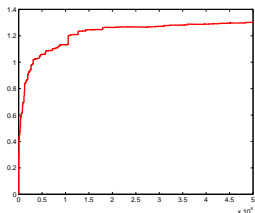


Choix de θ par Méthode 2

On calcule le drift comme la limite de

$$\theta_{n+1} = \theta_n - \gamma_{n+1} H_\alpha(\theta_n, Z_{n+1}) \quad H_\alpha(\theta, z) = \exp(-\alpha\sqrt{1+\theta^2}) \phi^2(z-\theta)(2t-z).$$

- 1 On trace l'évolution de la suite $\{\theta_n, n \geq 0\}$ qui converge vers $\theta_\star^{(2)} = 1.301$;
- 2 On compare les deux estimateurs (estimation + intervalles de confiance à 95%).



Coupler adaptation et simulation

Deux possibilités

① Approche 1:

- ① Mettre en oeuvre un algorithme de gradient stochastique pour trouver $\theta_{\star}^{(2)}$.
- ② Avec cette valeur du drift, calculer l'estimateur d'échantillonnage d'importance de la forme

$$n^{-1} \sum_{k=1}^n \tilde{\phi}(Z_k, \theta_{\star}^{(2)}).$$

② Approche 2: mettre en oeuvre l'algorithme suivant

- ① calculer $\tilde{\phi}(Z_{k+1}, \theta_k)$
- ② mettre à jour le drift: $\theta_{k+1} = \theta_k - \gamma_{k+1} H_{\alpha}(\theta_k, Z_{k+1})$

et calculer l'estimateur

$$n^{-1} \sum_{k=1}^n \tilde{\phi}(Z_{k+1}, \theta_k).$$

Attention: ce n'est pas une somme de v.a. i.i.d. donc l'étude de cet estimateur (convergence, normalité asymptotique, ...)

repose sur des outils de martingales

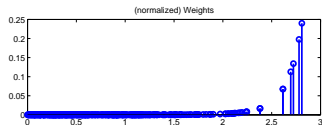
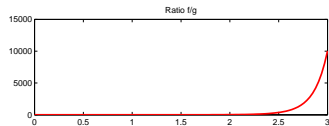
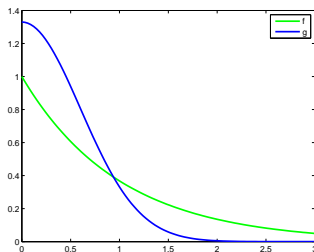
↪ nous reviendrons sur les **méthodes de simulation adaptatives** dans les prochains cours

Exemples de poids d'importance (I)

- La densité cible est $f(x) = \lambda \exp(-\lambda x) \mathbb{I}_{\mathbb{R}^+}(x)$.
- La densité instrumentale est $g(x) = \frac{2}{\sqrt{2\pi}\sigma} \exp(-0.5x^2/\sigma^2) \mathbb{I}_{\mathbb{R}^+}(x)$.

Le ratio d'importance est (sur \mathbb{R}^+)

$$\frac{f(x)}{g(x)} = \frac{\sqrt{2\pi}\lambda\sigma}{2} \exp(-\lambda x + 0.5x^2/\sigma^2) \longrightarrow +\infty \text{ quand } x \rightarrow \infty$$



gauche Densités f et g

droite (haut) $x \mapsto f(x)/g(x)$; (bas) calcul de 500 poids d'importance associés à 500 tirages sous g

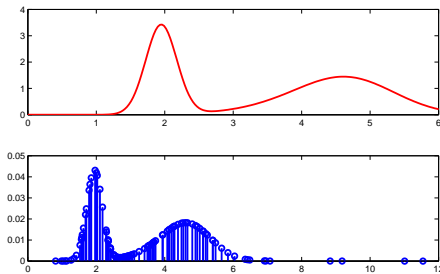
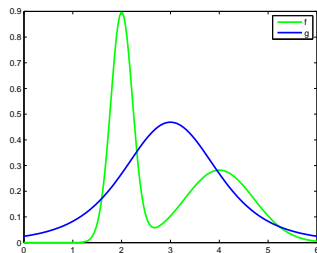
↪ seuls quelques points sont significatifs

► Suite

Exemples de poids d'importance (II)

- La densité cible est $f(x) = 0.5 \mathcal{N}(2, 0.05) + 0.5 \mathcal{N}(4, 0.5)$.
- La densité instrumentale est (Student 4 degrés de liberté, non centrée)
 $g(x) \propto (1 + (x - 3)^2/4)^{-5/2}$.

Le ratio d'importance est borné.



gauche Densités f et g

droite (haut) $x \mapsto f(x)/g(x)$; (bas) calcul de 100 poids d'importance associés à 100 tirages sous g

↔ Lois de Student conseillées pour le changement de loi (queues lourdes)