Réduction de la variance

Rappel :

X . var à valeurs dans $\mathbb{R}^d$ simulable

$H : \mathbb{R}^d \to \mathbb{R}$ application mesurable, connue explicitement tq $E(|H(X)|) < +\infty$.

But : Estimer $I = E(H(X))$ les outils $X_1, \ldots, X_m$ m-éch de X

Les outils : $X_1, \ldots, X_m$ m échantillons de X.

MMC : $\hat{I}_m = \frac{1}{m} \sum_{K=1}^{m} H(X_K)$ esb de I

$\hat{I}_m \xrightarrow[m \to +\infty]{\text{ps}} I$ (LFGN) fortement consistant

On suppose de plus $E((H(X))^2) < +\infty$, on a :

$$R_{\hat{I}_m}(I) = E\left((\hat{I}_m - I)^2\right) \underset{\hat{I}_m \text{ est sb de } I}{=} Var(\hat{I}_m) = \frac{Var(H(X))}{m}$$

$$\sqrt{m}\,(\hat{I}_m - I) \xrightarrow[m \to +\infty]{\text{Loi}} N(0, Var(H(X))) \quad \text{(TLC)}$$

$\tilde{S}_m^2 = \frac{1}{m-1} \sum_{K=1}^{m} (H(X_K) - \hat{I}_m)^2$ esb de $Var(H(X))$, de plus, $\tilde{S}_m^{(2)} \xrightarrow[m \to +\infty]{\text{ps}} Var(H(X))$

Lemme de Slutsky : $\underset{\text{fonction pivotale asymptotique pour } I}{Q_m(I)} = \frac{\sqrt{m}\,(\hat{I}_m - I)}{\sqrt{\tilde{S}_m^{(2)}}} \xrightarrow[m \to +\infty]{\text{Loi}} N(0,1)$

$\left[\hat{I}_m \pm \frac{\sqrt{\tilde{S}_m^{(2)}}}{\sqrt{m}}\, q_{1-\alpha/2}^{N(0,1)}\right]$ est un intervalle de confiance bilatéral symétrique de niveau asymptotique $1-\alpha$.

I/ Méthodes antithétiques : Variables antithétiques

Principe : Pour comparer deux méthodes, on doit utiliser la même échantillon $X_1, \ldots, X_m$ de X, càd : Trouver un nouveau estimateur $\tilde{I}_m$ (esb de I), fortement consistant tq $\boxed{Var(\tilde{I}_m) < Var(\hat{I}_m)}$

Exemples :

./ $X \hookrightarrow N(0,1)$ $I = E(H(X))$

Posons $A(u) = -u$, on constate que $A(X) = -X$ et X ont même loi

Posons $\tilde{I}_n = \frac{1}{n} \sum_{k=1}^{n} \frac{H(X_k) + H(A(X_k))}{2}$; $E(\tilde{I}_n) = I$

LFGN: $\tilde{I}_n \xrightarrow[n \to +\infty]{} I$ ps

$R_{\tilde{I}_n}(I) = Var(\tilde{I}_n) = \frac{1}{n} Var\left(\frac{H(X) + H(A(X))}{2}\right)$

Par suite, on doit comparer $Var\left(\frac{H(X)+H(A(X))}{2}\right)$ et $Var(H(X))$

or $Var\left(\frac{H(X)+H(A(X))}{2}\right) = \frac{1}{2}\left(Var(H(X)) + Cov(H(X), H(A(X))\right)$

Dès que $Cov(H(X), H(A(X))) \leq 0 \Rightarrow Var\left(\frac{H(X)+H(A(X))}{2}\right) \leq \frac{Var(H(X))}{2} \leq Var(H(X$

2/ $X \hookrightarrow \mathcal{U}(]0,1[)$ $I = E(H(X))$

Prenons $A(u) = 1 - u$, on constate $A(X) = 1 - X$ et X ont même loi

$\tilde{I}_n = \frac{1}{n} \sum_{k=1}^{n} \frac{H(X_k) + H(A(X_k))}{2}$

Pour les exemples 1 et 2, l'application : $u \mapsto A(u)$ est décroissante.

Sous l'hypothèse que H est monotone et A est décroissante

Montrons $cov(H(X), H(A(X))) \leq 0$

en effet; $u \mapsto H(u)$ et $u \mapsto H(A(u))$ sont de monotonie contraire

D'où $(H(u) - H(v))(H(A(u)) - H(A(v))) \leq 0$

Soient X_1 et X_2 deux v.ar indépendantes et de même loi que X

$(H(X_1) - H(X_2))(H(A(X_1)) - H(A(X_2))) \leq 0$

$\Rightarrow E((H(X_1) - H(X_2))(H(A(X_1)) - H(A(X_2))) \leq 0$

$\underset{\substack{X_1 \text{ et } X_2 \\ \text{indep}}}{=} E(H(X_1)H(A(X_2))) - E(\cancel{H(X_1) E(H(A(X_1))}) - E(H(X_1))E(H(A(X_2)))$
$+ E(\cancel{H(X_1) H(A(X_2))})$

$\underset{\substack{X_1 \text{ et } X_2 \\ \text{de même loi que } X}}{=} E(H(X)H(A(X))) - E(H(X)) E(H(A(X))) - E(H(X)) E(H(A(X)))$
$+ E(H(X).H(A(X)))$

$= 2\left(E(H(X).H(A(X))) - E(H(X)) E(H(A(X)))\right) = 2.cov(H(X), H(A(X)))$

d'où le résultat.

Proposition :

Hyp : • H est monotone

• A est décroissante tq $A(X)$ et X ont même loi ($\boxed{A(X)}$ s'appelle la variable antithétique)

Conc : $\mathrm{Cov}(H(X), H(A(X))) \leq 0$

Proposition :

$K(X) = H(A(X))$

Hyp : • H et K deux applications de $\boxed{\text{monotonie contraire}}$

• $H(X)$ et $K(X)$ ont même loi avec $E(H(X)^2) < +\infty$

Conc : $\boxed{\mathrm{Cov}(H(X), K(X)) \leq 0}$

MMC : $\hat{I}_n = \frac{1}{n} \sum_{k=1}^{n} H(X_k)$ $\qquad \tilde{I}_n = \frac{1}{n} \sum_{k=1}^{n} \frac{H(X_k) + K(X_k)}{2}$

$$\mathrm{Var}(\tilde{I}_n) = \frac{1}{4n} \mathrm{Var}(H(X) + K(X)) \leq \frac{1}{n} \mathrm{Var}(H(X)) = \mathrm{Var}(\hat{I}_n)$$

Généralisation :

On suppose que :

• $H : \mathbb{R}^d \longmapsto \mathbb{R}$ application mesurable croissante en chacune de ses coordonnées.

• $K : \mathbb{R}^d \longmapsto \mathbb{R}$ application mesurable décroissante en chacune de ses coordonnées

• $H(X)$ et $K(X)$ ont m̂ loi où X est une v.a à valeurs dans $\mathbb{R}^d$, simulable

MMC : $\hat{I}_n = \frac{1}{n} \sum_{k=1}^{n} H(X_k)$ $\qquad$ Méth. antithétique : $\boxed{\tilde{I}_n = \frac{1}{2n} \sum_{k=1}^{n} (H(X_k) + K(X_k))}$

$$\boxed{\mathrm{Var}(\tilde{I}_n) \leq \mathrm{Var}(\hat{I}_n)}$$

Preuve : voir TD3

II) Echantillonnage préférentiel : Fonction d'importance :

X va à valeurs dans $\mathbb{R}^d$ de densité f par rapport à la mesure de Lebesgue de $\mathbb{R}^d$. (X est simulable)

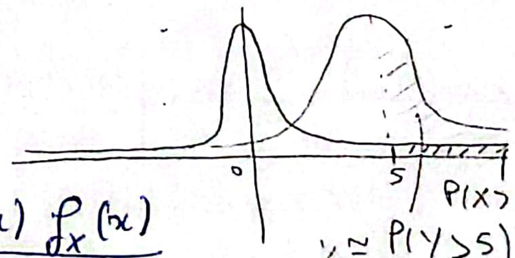
Exemples :

1/ $X \hookrightarrow N(0,1)$ $\qquad I = P(X>5) = E(1_{(X>5)}) = E(1_{]5;+\infty[}(X)) = E(H(X))$

MMC $\hat{I}_n = \sum_{k=1}^{n} \frac{1_{]5;+\infty[}(X_k)}{n} = \frac{1}{n} \sum_{k=1}^{n} H(X_k)$ $\qquad H(x) = 1_{]5;+\infty[}(x)$

En réalité $P(X>5) = 0{,}00\ldots$ (très proche de 0)

D'où $\hat{I}_n = \frac{1}{n} \sum_{k=1}^{n} H(X_k)$ (MMC)



$I = \int_{-\infty}^{+\infty} H(x) \cdot f_X(x)\, dx$ $\qquad f_X = f_{N(0,1)}$ $\rightarrow \frac{H(x) f_X(x)}{f_Y(x)}$

$= \int^{+\infty} \frac{H(x) f_X(x)}{} f_Y(x)\, dx$ $\qquad f_Y = f_{N(0,1)}$

$= E(K(Y))$ $\quad Y \hookrightarrow N(5,1)$ $\quad K(x) = \dfrac{H(x).f_X(x)}{f_Y(x)}$

$\tilde{I}_m = \frac{1}{m} \sum_{k=1}^{m} K(Y_k)$ où $Y_1, \dots, Y_m \hookrightarrow Y \hookrightarrow N(5,1)$

$Y_k = 5 + X_k$

$\mathbb{1}_{]5;+\infty[}(Y_k) \dfrac{f(Y_k)}{g(Y_k)}$) une chance sur 2 ($\neq 0$)

. X v.a à valeurs dans $\mathbb{R}^d$, de densité f, par rapport à la mesure de Lebesgue, simulable

. $H: \mathbb{R}^d \to \mathbb{R}$ mesurable, tq $E(|H(X)|) < +\infty$, qui est connue explicitement

But: Estimer $I = E(H(X))$, les outils : $X_1, \dots, X_m$ m éch de X

MCC: . $\hat{I}_m = \frac{1}{m} \sum_{k=1}^{m} H(X_k)$

. Z v.a à valeurs dans $\mathbb{R}^d$, de densité g par rapport à la mesure de Lebesgue.

Généralement, $Z = \varphi(X)$ $\quad Z_1, \dots, Z_m$ càd $Z_k = \varphi(X_k)$ $\quad X_i$ sont iid $\hookrightarrow X$

. hyp: $\overbrace{\text{supp}(H.f)}^{A} \subseteq \overbrace{\text{supp}(g)}^{B}$ $\quad \text{supp } K = \{x \in \mathbb{R}^d / K(x) \neq 0\}$

$I = E(H(X)) = \int_{\mathbb{R}^d} H(x).f(x)\,dx = \int_A H(x).f(x)\,dx = \int_A \dfrac{H(x).f(x)}{g(x)} g(x)\,dx$

$g(x) \neq 0$

$f \subset g$.

$= \int_B \dfrac{H(x).f(x)}{g(x)} g(x)\,dx = E\left(\dfrac{H(Z)f(Z)}{g(Z)}\right)$

$I = E(H(X)) = E\left(\dfrac{H(Z).f(Z)}{g(Z)}\right)$

$\tilde{I}_m = \frac{1}{m} \sum_{k=1}^{m} \dfrac{H(Z_k)f(Z_k)}{g(Z_k)}$ $\quad E(\tilde{I}_m) = E\left(\dfrac{H(Z)f(Z)}{g(Z)}\right) = I$

$\hookrightarrow$ est un esb de I

LFGN : $\tilde{I}_m \xrightarrow[m \to +\infty]{Pps} I$, càd, $\tilde{I}_m$ est fortement consistant

Sous l'hypothèse $E\left(\left|\dfrac{H(Z)f(Z)}{g(Z)}\right|^2\right) < +\infty$, TLC: $\sqrt{m}(\tilde{I}_m - I) \xrightarrow[m \to +\infty]{Loi} N\left(0, Var\left(\dfrac{H(Z)f(Z)}{g(Z)}\right)\right)$

$Var(\tilde{I}_m) = \frac{1}{m} Var\left(\dfrac{H(Z)f(Z)}{g(Z)}\right)$; $Var(\hat{I}_m) = \frac{1}{m} Var(H(X))$

$Var(H(X)) = E((H(X))^2) - (E(H(X)))^2 = E((H(X))^2) - I^2$

$Var\left(\dfrac{H(Z).f(Z)}{g(Z)}\right) = E\left(\left(\dfrac{H(Z)f(Z)}{g(Z)}\right)^2\right) - I^2$

But : Comparer $☺ = E(|H(X)|^2)$ $☺☺ = E\left(\left|\frac{H(Z).f(Z)}{g(Z)}\right|^2\right)$

Par suite, on a le problème suivant :

$$\min_{\substack{g \text{ est une densité} \\ \text{sur } \mathbb{R}^d,\ supp(H.f) \subseteq supp(g)}} E\left(\left|\frac{H(Z)f(Z)}{g(Z)}\right|^2\right)$$

Prenons $g^*(x) = \frac{1}{\alpha} f(x).|H(x)|$ avec $\alpha = \int_{\mathbb{R}^d} |H(x)| f(x)\,dx = E(|H(X)|)$

↳ Montrons que g^* est une solution du problème ci-dessus

$$☺☺^* = ☺☺_{g^*} = \cancel{\alpha^2}\, E\left(\left|\frac{H(Z)f(Z)}{g(Z)}\right|^2\right) = \alpha^2 = (E(|H(X)|))^2$$

$$☺☺_g = E\left(\frac{H(Z)f(Z)}{g(Z)}\right) = \int_{\mathbb{R}^d}\left(\frac{H(x)f(x)}{g(x)}\right)^2 g(x)\,dx = \int_{\mathbb{R}^d}\frac{(H(x))^2 f(x)}{g(x)}\, f(x)\,dx$$

$$☺☺_g = E\left(\frac{(H(X))^2 f(X)}{g(X)}\right) = E\left(\left(\frac{|H(Z)|\ f(Z)}{g(Z)}\right)^2\right) \geq \left(E\left(\frac{|H(Z)|f(Z)}{g(Z)}\right)\right)^2$$

$\varphi(x) = x^2$ est convexe

(*) inégalité de Jensen où bien $Var(T) \geq 0$ où $T = \frac{|H(Z)|f(Z)}{g(Z)}$

$$☺☺_g \geq \left(\int_{\mathbb{R}^d}\frac{|H(x)|f(x)}{g(x)}g(x)\,dx\right)^2 = \left(\int_{\mathbb{R}^d}|H(x)|f(x)\,dx\right)^2 = \alpha^2 = ☺☺_{g^*}$$

$supp\ g^* = supp(|H|.f)$

$$Var\left(\frac{H(Z).f(Z)}{g^*(Z)}\right) = E\left(\left|\frac{H(Z)f(Z)}{g^*(Z)}\right|^2\right) - I^2 = (E(|H(X)|))^2 - |E(H(X))|^2 = 0$$

dès que $H \geq 0$

Si $H \geq 0$:

alors $Var\left(\frac{H(Z)f(Z)}{g^*(Z)}\right) = 0 \iff \frac{H(x)f(x)}{g^*(x)} = cnste$

Le choix optimal : $\frac{H(x)f(x)}{g^*(x)} = cnst$

Exemple :

Méthode 1 : $I = P(X>5)$ $X \hookrightarrow N(0,1)$ $I = E(H(X))$

↳ $H(x) = 1_{]5;+\infty[}$

$Z = \theta + X \hookrightarrow N(\theta, 1)$ $\theta \in \mathbb{R}_+$

$X_K \longmapsto Z_K = \theta + X_K$ $f_{N(\theta,1)}(x) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-1}{2}(x-\theta)^2\right) = g(x)$

$$\tilde{I}_m = \frac{1}{m}\sum_{K=1}^{m}\frac{H(Z_K)f(Z_K)}{g(Z_K)}$$ $supp\{H.f_{N(0,1)}\} =]5;+\infty[\subseteq \mathbb{R} = supp\, g$

$$MCC : \hat{I}_m = \frac{1}{m}\sum_{K=1}^{m} H(X_K) = \frac{1}{m}\sum_{K=1}^{m}\frac{H(Z_K)f(Z_K)}{g(Z_K)}$$

→ $H(Z_K) = 1_{]5;+\infty[}$ (=

$= P(Z>5) = P(\theta+X>5$

$= P(X>5-\theta) = 1-$ ...

$L'(\theta) = f_{N(0,1)}(5-\theta) > 0$ $\quad f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$

θ	0	1/2	
$L(\theta)$			

$\theta = 5$ une chance sur 2 que la qte $\frac{H(Z).f(Z)}{g(Z)} \neq 0$

Méthode 2 : $Z = T+\theta$ $\quad T \hookrightarrow \mathcal{E}(1)$ $\quad I = P(X>5)$

$$E(K(Z)) = \int_0^{+\infty} K(t+\theta)\, e^{-t} dt = \int_\theta^{+\infty} K(z).\, e^{-(z-\theta)} dz = \int_{-\infty}^{+\infty} K(z).\, e^{-(z-\theta)} 1_{]\theta;+\infty[}(z).$$

$g = f_Z$ $\quad$ supp $g =]\theta;+\infty[$: supp $(H.f) \subset$ supp g

$]5;+\infty[\; = \;]\theta;+\infty[\Rightarrow \theta = 5$

$$\tilde{I}_m = \frac{1}{m}\sum_{K=1}^{m} \frac{H(Z_K)f(Z_K)}{g(Z_K)}$$

$$P(Z>5) = \int_5^{+\infty} e^{-(z-5)} 1_{]5;+\infty[}(z)\, dz = 1$$

La chance est certaine $(\neq 0)$

$\varphi = ?$ $\quad Z = \varphi(X)$ $\quad X \hookrightarrow N(0,1)$ la loi de Z s'appelle la loi exponentielle translatée

le but c'est de ne pas avoir 0 cad $\underset{=0}{\underline{Z_K < 5}}$ d'où $\neq 0$ pour prob $P(Z>5)$

III Méthode de variable de contrôle :

X v.a à valeurs dans $\mathbb{R}^d$, simulable

$H : \mathbb{R}^d \longmapsto \mathbb{R}$ connue explicitement tq $E(|H(X)|) < +\infty$

$\Rightarrow$ But : Estimer $I = E(H(X))$

Les outils : $X_1, \ldots, X_m$ m-échantillon de X

MMC : $\hat{I}_m = \frac{1}{m}\sum_{K=1}^{m} H(X_K)$ estimateur sans biais de I, fortement consistant

de plus $\sqrt{m}(\hat{I}_m - I) \xrightarrow[m\to+\infty]{Loi} N(0, Var(H(X)))$, dès que $E((H(X))^2) < +\infty$ (TLC

$Var(\hat{I}_m) = \frac{Var(H(X))}{m}$

Hyp : on suppose qu'il existe une application $H_0 : \mathbb{R}^d \to \mathbb{R}$ tq $E(H_0(X)) = m =$ connue explicitement

Posons $\hat{I}_m^b = \frac{1}{m}\sum_{K=1}^{m}\left(H(X_K) - b(H_0(X_K) - m)\right)$

$-E(\hat{I}_m^b) = E\left(H(X) - b(H_0(X) - m)\right) = E(H(X)) - b(\underbrace{E(H_0(X)) - m}_{=0}) = I$, cad

$\hat{I}_m^b$ est un esb de I

LFGN : $\hat{I}_m^b \xrightarrow[m\to+\infty]{Pps} E(H(X) - b(H_0(X) - m)) = I$

$\sqrt{m}(\hat{I}_m^b - I) \xrightarrow[m\to+\infty]{Loi} N\left(0, var\left(H(X) - b(H_0(X) - m)\right)\right)$

staff asra3 men MC $\Rightarrow$ تقرب نرسو لا va nulle

Dans la suite, on doit comparer $Var(H(X))$ et $Var(H(X) - b\, H_0(X))$

$$Var(H(X) - b H_0(X)) = Var(H(X)) + b^2 Var(H_0(X)) - 2b\, cov(H(X), H_0(X)) = L(b)$$

Pb : $\min_{b \in \mathbb{R}} L(b)$ $\quad L'(b) = 2b\, Var(H_0(X)) - 2 Cov(H(X), H_0(X))$

$$L'(b) = 0 \iff b^* = \frac{Cov(H(X), H_0(X))}{Var(H_0(X))} \Rightarrow \text{extremum}$$

$$L''(b) = 2 Var(H_0(X)) > 0 \Rightarrow b^* \text{ est minimum}$$

$$L(b^*) = Var(H(X)) - \frac{[Cov(H(X), H_0(X))]^2}{Var(H_0(X))} \leq Var(H(X))$$

Inconvénient : On doit connaît explicitement $b^* = \dfrac{E(H(X).H_0(X)) - E(H(X))E(H_0(X))}{Var(H_0(X))}$

Par suit, on doit estimer b^*, mais en gardant le m̂ échantillon $X_1, \dots, X_{l_m}, X_{l_m+1}, \dots$

$l_m \ll m$ mais $l_m \xrightarrow[m \to +\infty]{} +\infty$ et $\frac{l_m}{m} \xrightarrow[m \to +\infty]{} 0$ (Exempl : $l_m = [\sqrt{m}] \xrightarrow[m \to +\infty]{} +\infty$, $\frac{l_m}{m} \xrightarrow[m \to +\infty]{} 0$ $\quad m = 10000, l_m =$

$$b_m^* = \frac{\frac{1}{l_m} \sum_{k=1}^{l_m} H(X_k).H_0(X_k) - \frac{m}{l_m} \sum_{k=1}^{m} H(X_k)}{\frac{1}{l_m} \sum_{k=1}^{l_m} (H_0(X_k) - m)^2} = \frac{\sum_{k=1}^{l_m} H(X_k)(H_0(X_k) - m)}{\sum_{k=1}^{l_m} (H_0(X_k) - m)^2}$$

LFGN $\xrightarrow{pps}$ ($m \to +\infty$, $l_m \to +\infty$) : $E(H(X).H_0(X))$; $\quad n \to +\infty$, $l_n \to +\infty$ $\xrightarrow{pps}$ mI

$m \to +\infty, l_m \to +\infty$ $\xrightarrow{pps}$ $Var(H_0(X))$ $\qquad b^* \xleftarrow[+\infty]{m}$

$$\hat{I}_m^{b_m} = \frac{1}{m - l_m} \sum_{k=l_m+1}^{m} \left(H(X_k) - b_m^*(H_0(X_k) - m)\right)$$

$$E(\hat{I}_m^{b_m}) = \frac{1}{m - l_m} \sum_{k=l_m+1}^{m} E\left(H(X_k) - b_m^*(H_0(X_k) - m)\right) = I \quad \text{sans biais}$$

$\underbrace{E(H(X))}_{=0}$ $\quad \underbrace{E(b_m^*(H_0(X_k) - m))}_{=0}$

$\Rightarrow b_m^*$ s'exprime en fonction de $X_1, \dots, X_{l_m}$; $(X_1, \dots, X_{l_m}) \perp X_K$ pour $K > l_m$.

IV/ Methode de Stratification :

X v.a à valeurs dans $\mathbb{R}^d$ simulable

$H : \mathbb{R}^d \to \mathbb{R}$ connu explicitement

Z v.a à valeurs dans $\mathbb{R}^m$, telle que

$\to P(Z \in A_k) = P_k > 0$ (connu) $(k = 1, \dots, K)$ où A_k forment une partition de $\mathbb{R}^m$

$\to \mathcal{L}(X / Z \in A_k)$ est simulable pour $k = 1, \dots, K$

But : Estimer $I = E(H(X))$

TLC : $\hat{I}_m = \frac{1}{m} \sum_{k=1}^{m} H(X_k)$ e.s.b de I, fortement consistant

$Var(\hat{I}_m) = \frac{Var(H(X))}{m}$, avec la condition $E(H(X)^2) < +\infty$

$\xi = (B_k)_{k=1,\dots,K}$ forment une partition de Ω. Posons $\mathcal{B} = \sigma(\xi)$

$I = E(H(X)) = E(E(H(X)|\mathcal{B}))$

$E(H(X)/\mathcal{B}) = \sum_{k=1}^{K} \underbrace{E(H(X)|B_k)}_{\text{deterministe}} 1_{B_k} \Rightarrow I = \sum_{k=1}^{K} \overbrace{E(H(X)/B_k)}^{I^{(k)}} . p_k$

$= \sum_{k=1}^{K} I^{(k)} p_k$

où $p_k = P(B_k) = P(Z \in A_k) = E(1_{A_k}(Z))$

L'idée : Estimer $I^{(k)}$ par la MMC $(k=1,\dots,K)$

$I^{(k)} = E(H(X)/B_k) = E(H(X^{(k)}))$, où $\mathcal{L}(X^{(k)}) = \mathcal{L}(X/B_k)$ simulable

$\hat{I}_m^{(k)} = \frac{1}{m_k} \sum_{p=1}^{m_k} H(X_p^{(k)}) \qquad \sum_{k=1}^{K} m_k = m \qquad m_k \xrightarrow[m\to+\infty]{} +\infty$

$\hat{I}_m^{(k)}$ est esb de $I^{(k)}$, fortement consistant

! $Var(\hat{I}_m^{(k)}) = \frac{Var(H(X^{(k)}))}{m_k}$ Posons $\hat{I}_m^{strat} = \sum_{k=1}^{K} \hat{I}_m^{(k)} p_k$

$E(\hat{I}_m^{strat}) = \sum_{k=1}^{K} E(\hat{I}_m^{(k)}) p_k = \sum_{k=1}^{K} I^{(k)} p_k = I$ Cad $\hat{I}_k^{strat}$ est un esb de I

$\lim_{m\to+\infty} \hat{I}_m^{strat} = \sum_{k=1}^{K} (\lim \hat{I}_m^{(k)}) p_k = \sum_{k=1}^{K} I^{(k)} p_k = I$, cad $\hat{I}_m^{strat}$ est

1. fortement consistant.

MMC : $\hat{I}_m = \frac{1}{m} \sum_{\ell=1}^{m} H(X_\ell)$ esb de I, fortement consistant

$(X_\ell)_{\ell=1,\dots,m}$ sont iid $\hookrightarrow X$

$Var(\hat{I}_m) = \frac{Var(H(X))}{m}$, avec la condition $E((H(X))^2) < +\infty$

1 $Var(\hat{I}_m^{strat}) = \sum_{k=1}^{K} Var(I_m^{(k)}) p_k^2 = \sum_{k=1}^{K} \frac{Var(H(X^{(k)}))}{m_k} p_k^2$

$= \frac{1}{m} \sum_{k=1}^{K} \frac{Var(H(X^{(k)}))}{\frac{m_k}{m}} p_k^2$

$= \frac{1}{m} \sum_{k=1}^{K} \frac{p_k^2}{q_k} Var(H(X^{(k)}))$ avec $q_k = \frac{m_k}{m}$

But : Définir les m_k $(k=1,\dots,K) \Leftrightarrow$ Définir les q_k $(k=1,\dots,K)$

et vérifier que $\underbrace{\sum_{k=1}^{K} \frac{p_k^2}{q_k} Var(H(X^{(k)}))}_{(*)} \leq Var(H(X))$

Cas particulier : Allocation proportionnelle $q_k = p_k$ $(k=1,\dots,K)$

$(*) = \sum_{k=1}^{K} Var(H(X^{(k)})) p_k$

$Var(H(X)) = E((H(X))^2) - (E(H(X)))^2 = E(\underbrace{E((H(X))^2/\mathcal{B})}_{(**)}) - (\sum_{k=1}^{K} I^{(k)} p_k)$

avec $I^{(k)} = E(H(X)/B_k) = E(H(X^{(k)}))$

$(**) = \sum_{k=1}^{K} E((H(X))^2/B_k) 1_{B_k}$

$$\mathrm{Var}(H(X)) = \sum_{k=1}^{K} E\left((H(X^{(k)}))^2\right) p_k - \sum_{k=1}^{K} \left(E(H(X^{(k)}))^2\right) p_k +$$

$$\left[\sum_{k=1}^{K} \left(E(H(X^{(k)}))\right)^2 p_k - \left(\sum_{k=1}^{K} E(H(X^{(k)})) p_k\right)^2\right]$$

$$= \underbrace{\sum_{k=1}^{K} \mathrm{Var}(H(X^{(k)})) p_k}_{⊛} + [\,]$$

Pour finir, on doit montrer que $[\,] \geq 0$.

f convexe $\quad \lambda_k \geq 0$ et $\sum_{k=1}^{K} \lambda_k = 1 \quad f\left(\sum_{k=1}^{K} \lambda_k x_k\right) \leq \sum_{k=1}^{K} \lambda_k f(x_k)$

$f(x) = x^2$

$\lambda_k = p_k$

$x_k = E(H(X^{(k)}))$

$$\left(\sum_{k=1}^{K} p_k E(H(X^{(k)}))\right)^2 \leq \sum_{k=1}^{K} p_k \left(E(H(X^{(k)}))\right)^2 \Rightarrow [\,] \geq 0$$

Allocation optimale: Pb : $\min\limits_{\substack{q_k > 0 \\ \sum q_k = 1}} \sum_{k=1}^{K} \frac{p_k^2}{q_k} \mathrm{Var}(H(X^{(k)}))$

Thm de KTT :

$$L(\lambda, q_1, \ldots, q_K) = \sum_{k=1}^{K} \frac{p_k^2}{q_k} \mathrm{Var}(H(X^{(k)})) + \lambda \left(\sum_{k=1}^{K} q_k - 1\right)$$

$$\nabla L = 0_{\mathbb{R}^{1+K}} \Leftrightarrow \begin{cases} \sum_{k=1}^{K} q_k = 1 \\ \frac{dL}{dq_\ell} = -\frac{p_\ell^2}{q_\ell^2} \mathrm{Var}(H(X^{(\ell)})) + \lambda \Leftrightarrow q_\ell = \frac{p_\ell}{\sqrt{\lambda}} \cdot \sqrt{\mathrm{Var}(H(X^{(\ell)}))} \end{cases}$$

$$1 = \sum_{\ell=1}^{K} q_\ell = \frac{1}{\sqrt{\lambda}} \cdot \sum_{\ell=1}^{K} p_\ell \sqrt{\mathrm{Var}(H(X^{(\ell)}))} \Rightarrow \lambda = \left(\sum_{k=1}^{K} p_k \sqrt{\mathrm{Var}(H(X^{(k)}))}\right)^2$$

$$q_\ell = \frac{p_\ell \sqrt{\mathrm{Var}(H(X^{(\ell)}))}}{\sum_{k=1}^{K} p_k \sqrt{\mathrm{Var}(H(X^{(k)}))}} \qquad \ell = 1, \ldots, K$$

Scanné avec CamScanner