Variables de contrôle (VC)

Idée: exploiter de l'information auxiliaire pour faire une correction à l'estimateur.

Variables de contrôle (VC)

Idée: exploiter de l'information auxiliaire pour faire une correction à l'estimateur.

Commençons par le cas des VC linéaires.

Variables de contrôle (VC)

Idée: exploiter de l'information auxiliaire pour faire une correction à l'estimateur.

Commençons par le cas des VC linéaires.

Soit X un estimateur sans biais de μ et $\mathbf{C} = (C^{(1)}, \ldots, C^{(q)})^{\mathsf{t}}$ des VCs corrélées avec X, d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \ldots, \nu^{(q)})^{\mathsf{t}}$.

Variables de contrôle (VC)

Idée: exploiter de l'information auxiliaire pour faire une correction à l'estimateur.

Commençons par le cas des VC linéaires.

Soit X un estimateur sans biais de μ et $\mathbf{C} = (C^{(1)}, \ldots, C^{(q)})^{\mathsf{t}}$ des VCs corrélées avec X, d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \ldots, \nu^{(q)})^{\mathsf{t}}$.

L'estimateur avec VC est:

$$X_{\mathrm{c}} = X - \boldsymbol{\beta}^{\mathsf{t}}(\mathbf{C} - \boldsymbol{\nu}) = X - \sum_{\ell=1}^{q} \beta_\ell (C^{(\ell)} - \nu^{(\ell)}),$$

où $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)^{\mathsf{t}}$ (des constantes).
On a $\mathbb{E}[X_{\mathrm{c}}] = \mathbb{E}[X] = \mu$.

Variables de contrôle (VC)

Idée: exploiter de l'information auxiliaire pour faire une correction à l'estimateur.

Commençons par le cas des VC linéaires.

Soit X un estimateur sans biais de μ et $\mathbf{C} = (C^{(1)}, \dots, C^{(q)})^{\mathsf{t}}$ des VCs corrélées avec X, d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^{\mathsf{t}}$.

L'estimateur avec VC est:

$$X_{\mathrm{c}} = X - \boldsymbol{\beta}^{\mathsf{t}}(\mathbf{C} - \boldsymbol{\nu}) = X - \sum_{\ell=1}^{q} \beta_\ell (C^{(\ell)} - \nu^{(\ell)}),$$

où $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^{\mathsf{t}}$ (des constantes).
On a $\mathbb{E}[X_{\mathrm{c}}] = \mathbb{E}[X] = \mu$.

Comment choisir $\boldsymbol{\beta}$?

Variables de contrôle (VC)

Idée: exploiter de l'information auxiliaire pour faire une correction à l'estimateur.

Commençons par le cas des VC linéaires.

Soit X un estimateur sans biais de μ et $\mathbf{C} = (C^{(1)}, \ldots, C^{(q)})^{\mathsf{t}}$ des VCs corrélées avec X, d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \ldots, \nu^{(q)})^{\mathsf{t}}$.

L'estimateur avec VC est:

$$X_{\mathrm{c}} = X - \boldsymbol{\beta}^{\mathsf{t}}(\mathbf{C} - \boldsymbol{\nu}) = X - \sum_{\ell=1}^{q} \beta_\ell (C^{(\ell)} - \nu^{(\ell)}),$$

où $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)^{\mathsf{t}}$ (des constantes).
On a $\mathbb{E}[X_{\mathrm{c}}] = \mathbb{E}[X] = \mu$.

Comment choisir $\boldsymbol{\beta}$?

Soient $\boldsymbol{\Sigma}_{\mathrm{C}} = \mathrm{Cov}[\mathbf{C}]$ et $\boldsymbol{\Sigma}_{\mathrm{CX}} = (\mathrm{Cov}(X, C^{(1)}), \ldots, \mathrm{Cov}(X, C^{(q)}))^{\mathsf{t}}$.

Variables de contrôle (VC)

Idée: exploiter de l'information auxiliaire pour faire une correction à l'estimateur.

Commençons par le cas des VC linéaires.

Soit X un estimateur sans biais de μ et $\mathbf{C} = (C^{(1)}, \ldots, C^{(q)})^{\mathsf{t}}$ des VCs corrélées avec X, d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \ldots, \nu^{(q)})^{\mathsf{t}}$.

L'estimateur avec VC est:

$$X_{\mathrm{c}} = X - \boldsymbol{\beta}^{\mathsf{t}}(\mathbf{C} - \boldsymbol{\nu}) = X - \sum_{\ell=1}^{q} \beta_\ell (C^{(\ell)} - \nu^{(\ell)}),$$

où $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)^{\mathsf{t}}$ (des constantes).
On a $\mathbb{E}[X_{\mathrm{c}}] = \mathbb{E}[X] = \mu$.

Comment choisir $\boldsymbol{\beta}$?

Soient $\boldsymbol{\Sigma}_{\mathrm{C}} = \mathrm{Cov}[\mathbf{C}]$ et $\boldsymbol{\Sigma}_{\mathrm{CX}} = (\mathrm{Cov}(X, C^{(1)}), \ldots, \mathrm{Cov}(X, C^{(q)}))^{\mathsf{t}}$.

Hypothèse VC1: $\mathrm{Var}[X] = \sigma^2 < \infty$, $\boldsymbol{\Sigma}_{\mathrm{C}}$ et $\boldsymbol{\Sigma}_{\mathrm{CX}}$ sont finies, et $\boldsymbol{\Sigma}_{\mathrm{C}}$ est définie positive (et donc inversible).

On a alors

$$\mathrm{Var}[X_{\mathrm{c}}] = \mathrm{Var}[X] + \boldsymbol{\beta}^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{C}} \boldsymbol{\beta} - 2\boldsymbol{\beta}^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{CX}}.$$

On a alors

$$\mathrm{Var}[X_\mathrm{c}] = \mathrm{Var}[X] + \boldsymbol{\beta}^\mathsf{t}\boldsymbol{\Sigma}_\mathrm{C}\boldsymbol{\beta} - 2\boldsymbol{\beta}^\mathsf{t}\boldsymbol{\Sigma}_\mathrm{CX}.$$

Pour minimiser par rapport à $\boldsymbol{\beta}$, on met le gradient par rapport à $\boldsymbol{\beta}$ égal à zéro:

$$0 = \nabla_{\boldsymbol{\beta}}\mathrm{Var}[X_\mathrm{c}] = 2\boldsymbol{\Sigma}_\mathrm{C}\boldsymbol{\beta} - 2\boldsymbol{\Sigma}_\mathrm{CX}.$$

On a alors

$$\mathrm{Var}[X_{\mathrm{c}}] = \mathrm{Var}[X] + \boldsymbol{\beta}^{\mathsf{t}}\boldsymbol{\Sigma}_{\mathrm{C}}\boldsymbol{\beta} - 2\boldsymbol{\beta}^{\mathsf{t}}\boldsymbol{\Sigma}_{\mathrm{CX}}.$$

Pour minimiser par rapport à $\boldsymbol{\beta}$, on met le gradient par rapport à $\boldsymbol{\beta}$ égal à zéro:

$$0 = \nabla_{\boldsymbol{\beta}}\mathrm{Var}[X_{\mathrm{c}}] = 2\boldsymbol{\Sigma}_{\mathrm{C}}\boldsymbol{\beta} - 2\boldsymbol{\Sigma}_{\mathrm{CX}}.$$

Le minimum est donc atteint pour

$$\boldsymbol{\beta} = \boldsymbol{\beta}^{*} = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}},$$

On a alors

$$\mathrm{Var}[X_{\mathrm{c}}] = \mathrm{Var}[X] + \boldsymbol{\beta}^{\mathrm{t}}\boldsymbol{\Sigma}_{\mathrm{C}}\boldsymbol{\beta} - 2\boldsymbol{\beta}^{\mathrm{t}}\boldsymbol{\Sigma}_{\mathrm{CX}}.$$

Pour minimiser par rapport à $\boldsymbol{\beta}$, on met le gradient par rapport à $\boldsymbol{\beta}$ égal à zéro:

$$0 = \nabla_{\boldsymbol{\beta}}\mathrm{Var}[X_{\mathrm{c}}] = 2\boldsymbol{\Sigma}_{\mathrm{C}}\boldsymbol{\beta} - 2\boldsymbol{\Sigma}_{\mathrm{CX}}.$$

Le minimum est donc atteint pour

$$\boldsymbol{\beta} = \boldsymbol{\beta}^{*} = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}},$$

qui donne la variance minimale

$$\mathrm{Var}[X_{\mathrm{c}}] \;=\; (1 - R^2_{\mathrm{CX}})\mathrm{Var}[X] \stackrel{\mathrm{def}}{=} \sigma^2_{\mathrm{c}},$$

où

$$R^2_{\mathrm{CX}} = \frac{\boldsymbol{\Sigma}^{\mathrm{t}}_{\mathrm{CX}}\boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}}{\mathrm{Var}[X]}$$

(le carré du coefficient de corrélation multiple entre $\mathbf{C}$ et X)
et la variance est réduite par le facteur $1 - R^2_{\mathrm{CX}} = \sigma^2_{\mathrm{c}}/\sigma^2$.

On a alors

$$\mathrm{Var}[X_\mathrm{c}] = \mathrm{Var}[X] + \boldsymbol{\beta}^\mathsf{t}\boldsymbol{\Sigma}_\mathrm{C}\boldsymbol{\beta} - 2\boldsymbol{\beta}^\mathsf{t}\boldsymbol{\Sigma}_\mathrm{CX}.$$

Pour minimiser par rapport à $\boldsymbol{\beta}$, on met le gradient par rapport à $\boldsymbol{\beta}$ égal à zéro:

$$0 = \nabla_{\boldsymbol{\beta}}\mathrm{Var}[X_\mathrm{c}] = 2\boldsymbol{\Sigma}_\mathrm{C}\boldsymbol{\beta} - 2\boldsymbol{\Sigma}_\mathrm{CX}.$$

Le minimum est donc atteint pour

$$\boldsymbol{\beta} = \boldsymbol{\beta}^* = \boldsymbol{\Sigma}_\mathrm{C}^{-1}\boldsymbol{\Sigma}_\mathrm{CX},$$

qui donne la variance minimale

$$\mathrm{Var}[X_\mathrm{c}] \;=\; (1 - R^2_\mathrm{CX})\mathrm{Var}[X] \stackrel{\mathrm{def}}{=} \sigma_\mathrm{c}^2,$$

où

$$R^2_\mathrm{CX} = \frac{\boldsymbol{\Sigma}^\mathsf{t}_\mathrm{CX}\boldsymbol{\Sigma}_\mathrm{C}^{-1}\boldsymbol{\Sigma}_\mathrm{CX}}{\mathrm{Var}[X]}$$

(le carré du coefficient de corrélation multiple entre $\mathbf{C}$ et X)
et la variance est réduite par le facteur $1 - R^2_\mathrm{CX} = \sigma_\mathrm{c}^2/\sigma^2$.

Mais avec $\boldsymbol{\beta} \neq \boldsymbol{\beta}^*$, la variance peut augmenter.

Types de VCs

(a) variables internes, basées sur des quantités déjà calculées durant la simulation;

Types de VCs

(a) variables internes, basées sur des quantités déjà calculées durant la simulation;

(b) variables externes, obtenues par des simulations additionnelles;

Types de VCs

(a) variables internes, basées sur des quantités déjà calculées durant la simulation;

(b) variables externes, obtenues par des simulations additionnelles;

(c) VCs implicites obtenues via une moyenne pondérée.
Soient $X^{(0)}, \ldots, X^{(q)}$ des estimateurs sans biais de μ. Posons

$$X_{\mathrm{c}} = \sum_{\ell=0}^{q} \beta_\ell X^{(\ell)} = X^{(0)} - \sum_{\ell=1}^{q} \beta_k (X^{(0)} - X^{(\ell)})$$

où $\sum_{\ell=0}^{q} \beta_\ell = 1$.

Types de VCs

(a) variables internes, basées sur des quantités déjà calculées durant la simulation;

(b) variables externes, obtenues par des simulations additionnelles;

(c) VCs implicites obtenues via une moyenne pondérée.
Soient $X^{(0)}, \ldots, X^{(q)}$ des estimateurs sans biais de μ. Posons

$$X_{\mathrm{c}} = \sum_{\ell=0}^{q} \beta_\ell X^{(\ell)} = X^{(0)} - \sum_{\ell=1}^{q} \beta_k (X^{(0)} - X^{(\ell)})$$

où $\sum_{\ell=0}^{q} \beta_\ell = 1$.
On peut interpréter $C^{(\ell)} = X^{(0)} - X^{(\ell)}$, $\ell = 1, \ldots, q$, comme VC pour $X = X^{(0)}$.

Estimation de β^*: propriétés asymptotiques

En pratique, on ne connait pas $\boldsymbol{\beta}^* = \mathbf{\Sigma}_{\mathrm{C}}^{-1}\mathbf{\Sigma}_{\mathrm{CX}}$
(parfois $\mathbf{\Sigma}_{\mathrm{C}}$, mais jamais $\mathbf{\Sigma}_{\mathrm{CX}}$).

Estimation de β^*: propriétés asymptotiques

En pratique, on ne connait pas $\boldsymbol{\beta}^* = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}$
(parfois $\boldsymbol{\Sigma}_{\mathrm{C}}$, mais jamais $\boldsymbol{\Sigma}_{\mathrm{CX}}$).

On peut l'estimer, disons par $\hat{\boldsymbol{\beta}}_n$, calculé à partir de $(X_1, \mathbf{C}_1), \ldots, (X_n, \mathbf{C}_n)$.

Estimation de β^*: propriétés asymptotiques

En pratique, on ne connait pas $\boldsymbol{\beta}^* = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}$
(parfois $\boldsymbol{\Sigma}_{\mathrm{C}}$, mais jamais $\boldsymbol{\Sigma}_{\mathrm{CX}}$).

On peut l'estimer, disons par $\hat{\boldsymbol{\beta}}_n$, calculé à partir de $(X_1, \mathbf{C}_1), \ldots, (X_n, \mathbf{C}_n)$.
Posons

$$X_{\mathrm{ce},i} = X_i - \hat{\boldsymbol{\beta}}_n^{\mathsf{t}}(\mathbf{C}_i - \boldsymbol{\nu})$$

et

$$\bar{X}_{\mathrm{ce},n} = \bar{X}_n - \hat{\boldsymbol{\beta}}_n^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

Estimation de β^*: propriétés asymptotiques

En pratique, on ne connait pas $\boldsymbol{\beta}^* = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}$
(parfois $\boldsymbol{\Sigma}_{\mathrm{C}}$, mais jamais $\boldsymbol{\Sigma}_{\mathrm{CX}}$).

On peut l'estimer, disons par $\hat{\boldsymbol{\beta}}_n$, calculé à partir de $(X_1, \mathbf{C}_1), \ldots, (X_n, \mathbf{C}_n)$.
Posons

$$X_{\mathrm{ce},i} = X_i - \hat{\boldsymbol{\beta}}_n^{\mathsf{t}}(\mathbf{C}_i - \boldsymbol{\nu})$$

et

$$\bar{X}_{\mathrm{ce},n} = \bar{X}_n - \hat{\boldsymbol{\beta}}_n^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

Théorème. Sous l'hypothèse CV1, lorsque $n \to \infty$, si $\hat{\boldsymbol{\beta}}_n \Rightarrow \boldsymbol{\beta}^*$, alors

$$\begin{aligned}
\sqrt{n}(\bar{X}_{\mathrm{c},n} - \bar{X}_{\mathrm{ce},n}) &\Rightarrow 0, \\
S^2_{\mathrm{ce},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n}(X_{\mathrm{ce},i} - \bar{X}_{\mathrm{ce},n})^2 &\Rightarrow \sigma_{\mathrm{c}}^2, \\
\frac{\sqrt{n}(\bar{X}_{\mathrm{ce},n} - \mu)}{S_{\mathrm{ce},n}} \Rightarrow \frac{\sqrt{n}(\bar{X}_{\mathrm{c},n} - \mu)}{\sigma_{\mathrm{c}}} &\Rightarrow N(0,1).
\end{aligned}$$

On peut utiliser ce théorème pour calculer un IC pour μ, en supposant que $\sqrt{n}(\bar{X}_{\mathrm{ce},n} - \mu)/S_{\mathrm{ce},n} \sim N(0,1)$.

On peut utiliser ce théorème pour calculer un IC pour μ, en supposant que $\sqrt{n}(\bar{X}_{\mathrm{ce},n} - \mu)/S_{\mathrm{ce},n} \sim N(0,1)$.

Comment construire $\hat{\boldsymbol{\beta}}_n$?

On peut utiliser ce théorème pour calculer un IC pour μ, en supposant que $\sqrt{n}(\bar{X}_{\mathrm{ce},n} - \mu)/S_{\mathrm{ce},n} \sim N(0,1)$.

Comment construire $\hat{\boldsymbol{\beta}}_n$? Méthode de base:

$$\hat{\boldsymbol{\beta}}_n = \hat{\boldsymbol{\Sigma}}_{\mathrm{C}}^{-1} \hat{\boldsymbol{\Sigma}}_{\mathrm{CX}}$$

où les éléments de $\hat{\boldsymbol{\Sigma}}_{\mathrm{C}}$ et $\hat{\boldsymbol{\Sigma}}_{\mathrm{CX}}$ sont

$$\begin{aligned}
\hat{\sigma}_{\mathrm{C}}^{(\ell,k)} &= \frac{1}{n-1} \sum_{i=1}^{n} (C_i^{(\ell)} - \bar{C}_n^{(\ell)})(C_i^{(k)} - \bar{C}_n^{(k)}), \\
\hat{\sigma}_{\mathrm{CX}}^{(\ell)} &= \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)(C_i^{(\ell)} - \bar{C}_n^{(\ell)}).
\end{aligned}$$

On peut utiliser ce théorème pour calculer un IC pour μ, en supposant que $\sqrt{n}(\bar{X}_{\mathrm{ce},n} - \mu)/S_{\mathrm{ce},n} \sim N(0,1)$.

Comment construire $\hat{\boldsymbol{\beta}}_n$? Méthode de base:

$$\hat{\boldsymbol{\beta}}_n = \hat{\boldsymbol{\Sigma}}_{\mathrm{C}}^{-1} \hat{\boldsymbol{\Sigma}}_{\mathrm{CX}}$$

où les éléments de $\hat{\boldsymbol{\Sigma}}_{\mathrm{C}}$ et $\hat{\boldsymbol{\Sigma}}_{\mathrm{CX}}$ sont

$$\begin{aligned}
\hat{\sigma}_{\mathrm{C}}^{(\ell,k)} &= \frac{1}{n-1}\sum_{i=1}^{n}(C_i^{(\ell)} - \bar{C}_n^{(\ell)})(C_i^{(k)} - \bar{C}_n^{(k)}), \\
\hat{\sigma}_{\mathrm{CX}}^{(\ell)} &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)(C_i^{(\ell)} - \bar{C}_n^{(\ell)}).
\end{aligned}$$

Variantes: remplacer $\bar{C}_n^{(\ell)}$ par $\nu^{(\ell)}$, ou encore $\hat{\sigma}_{\mathrm{C}}^{(\ell,k)}$ par $\sigma_{\mathrm{C}}^{(\ell,k)}$, si connu.

Cas multinormal

Dans le cas où $\begin{pmatrix} X_i \\ \mathbf{C}_i \end{pmatrix} \sim$ normal, on peut utiliser la théorie de la régression linéaire (avec estimateurs moindres carrés) pour le modèle

$$X = \mu + \boldsymbol{\beta}^{\mathsf{t}}(\mathbf{C} - \boldsymbol{\nu}) + \epsilon$$

où $\epsilon \sim N(0, \sigma_\epsilon^2)$.

Cas multinormal

Dans le cas où $\begin{pmatrix} X_i \\ \mathbf{C}_i \end{pmatrix} \sim$ normal, on peut utiliser la théorie de la régression linéaire (avec estimateurs moindres carrés) pour le modèle

$$X = \mu + \boldsymbol{\beta}^{\mathsf{t}}(\mathbf{C} - \boldsymbol{\nu}) + \epsilon$$

où $\epsilon \sim N(0, \sigma_\epsilon^2)$. Si on définit

$$\tilde{S}^2_{\mathrm{ce},n} = \frac{n}{n-q-1}\left(\frac{1}{n} + \frac{(\bar{\mathbf{C}}_n - \boldsymbol{\nu})^{\mathsf{t}}\hat{\boldsymbol{\Sigma}}_{\mathrm{C}}^{-1}(\bar{\mathbf{C}}_n - \boldsymbol{\nu})}{n-1}\right)\sum_{i=1}^{n}(X_{\mathrm{ce},i} - \bar{X}_{\mathrm{ce},n})^2,$$

où $\hat{\boldsymbol{\beta}}_n$ utilise les covariances empiriques, on a

Théorème. Si les $(X_i, \mathbf{C}_i^{\mathsf{t}})^{\mathsf{t}}$ sont i.i.d normaux, alors

$$\begin{aligned}
\mathbb{E}[\bar{X}_{\mathrm{ce},n}] &= \mu \qquad \text{(aucun biais)}, \\
\mathbb{E}[\tilde{S}^2_{\mathrm{ce},n}/n] &= \mathrm{Var}[\bar{X}_{\mathrm{ce},n}] \;=\; \frac{n-2}{n-q-2}(1-R^2_{\mathrm{CX}})\mathrm{Var}[\bar{X}_n], \\
\sqrt{n}(\bar{X}_{\mathrm{ce},n}-\mu)/\tilde{S}_{\mathrm{ce},n} &\sim \text{Student-t}(n-q-1) \qquad \text{(loi exacte)}.
\end{aligned}$$

Permet de calculer un IC avec couverture exacte pour n fini.

Théorème. Si les $(X_i, \mathbf{C}_i^{\mathsf{t}})^{\mathsf{t}}$ sont i.i.d normaux, alors

$$\begin{aligned}
\mathbb{E}[\bar{X}_{\mathrm{ce},n}] &= \mu \qquad \text{(aucun biais)}, \\
\mathbb{E}[\tilde{S}^2_{\mathrm{ce},n}/n] &= \mathrm{Var}[\bar{X}_{\mathrm{ce},n}] = \frac{n-2}{n-q-2}(1-R^2_{\mathrm{CX}})\mathrm{Var}[\bar{X}_n], \\
\sqrt{n}(\bar{X}_{\mathrm{ce},n}-\mu)/\tilde{S}_{\mathrm{ce},n} &\sim \mathrm{Student\text{-}t}(n-q-1) \qquad \text{(loi exacte)}.
\end{aligned}$$

Permet de calculer un IC avec couverture exacte pour n fini.

Facteur d'inflation de la variance $(n-2)/(n-q-2) > 1$ dû à l'estimation de $\boldsymbol{\beta}^*$.

Théorème. Si les $(X_i, \mathbf{C}_i^{\mathsf{t}})^{\mathsf{t}}$ sont i.i.d normaux, alors

$$\begin{aligned}
\mathbb{E}[\bar{X}_{\mathrm{ce},n}] &= \mu \qquad \text{(aucun biais)},\\
\mathbb{E}[\tilde{S}^2_{\mathrm{ce},n}/n] &= \mathrm{Var}[\bar{X}_{\mathrm{ce},n}] = \frac{n-2}{n-q-2}(1-R^2_{\mathrm{CX}})\mathrm{Var}[\bar{X}_n],\\
\sqrt{n}(\bar{X}_{\mathrm{ce},n}-\mu)/\tilde{S}_{\mathrm{ce},n} &\sim \mathrm{Student\text{-}t}(n-q-1) \qquad \text{(loi exacte)}.
\end{aligned}$$

Permet de calculer un IC avec couverture exacte pour n fini.

Facteur d'inflation de la variance $(n-2)/(n-q-2) > 1$ dû à l'estimation de $\boldsymbol{\beta}^*$.

Si on a déjà q VC, l'ajout d'une nouvelle VC n'est rentable que si la valeur de $(1-R^2_{\mathrm{CX}})$ est réduite d'une fraction $\geq 1/(n-q-2)$.

Théorème. Si les $(X_i, \mathbf{C}_i^{\mathsf{t}})^{\mathsf{t}}$ sont i.i.d normaux, alors

$$\begin{aligned}
\mathbb{E}[\bar{X}_{\mathrm{ce},n}] &= \mu \qquad \text{(aucun biais)}, \\
\mathbb{E}[\tilde{S}^2_{\mathrm{ce},n}/n] &= \mathrm{Var}[\bar{X}_{\mathrm{ce},n}] = \frac{n-2}{n-q-2}(1-R^2_{\mathrm{CX}})\mathrm{Var}[\bar{X}_n], \\
\sqrt{n}(\bar{X}_{\mathrm{ce},n}-\mu)/\tilde{S}_{\mathrm{ce},n} &\sim \mathrm{Student\text{-}t}(n-q-1) \qquad \text{(loi exacte)}.
\end{aligned}$$

Permet de calculer un IC avec couverture exacte pour n fini.

Facteur d'inflation de la variance $(n-2)/(n-q-2) > 1$ dû à l'estimation de $\boldsymbol{\beta}^*$.

Si on a déjà q VC, l'ajout d'une nouvelle VC n'est rentable que si la valeur de $(1-R^2_{\mathrm{CX}})$ est réduite d'une fraction $\geq 1/(n-q-2)$.

Si n est petit et que la loi des $\begin{pmatrix} X_i \\ \mathbf{C}_i \end{pmatrix}$ est loin de la loi normale, il faut trouver des manières de contrôler le biais. Voir les notes (splitting).

Théorème. Si les $(X_i, \mathbf{C}_i^{\mathsf{t}})^{\mathsf{t}}$ sont i.i.d normaux, alors

$$
\begin{aligned}
\mathbb{E}[\bar{X}_{\mathrm{ce},n}] &= \mu \qquad \text{(aucun biais)},\\
\mathbb{E}[\tilde{S}^2_{\mathrm{ce},n}/n] &= \mathrm{Var}[\bar{X}_{\mathrm{ce},n}] = \frac{n-2}{n-q-2}(1-R^2_{\mathrm{CX}})\mathrm{Var}[\bar{X}_n],\\
\sqrt{n}(\bar{X}_{\mathrm{ce},n}-\mu)/\tilde{S}_{\mathrm{ce},n} &\sim \mathrm{Student\text{-}t}(n-q-1) \qquad \text{(loi exacte)}.
\end{aligned}
$$

Permet de calculer un IC avec couverture exacte pour n fini.

Facteur d'inflation de la variance $(n-2)/(n-q-2) > 1$ dû à l'estimation de $\boldsymbol{\beta}^*$.

Si on a déjà q VC, l'ajout d'une nouvelle VC n'est rentable que si la valeur de $(1-R^2_{\mathrm{CX}})$ est réduite d'une fraction $\geq 1/(n-q-2)$.

Si n est petit et que la loi des $\begin{pmatrix} X_i \\ \mathbf{C}_i \end{pmatrix}$ est loin de la loi normale, il faut trouver des manières de contrôler le biais. Voir les notes (splitting).

On peut aussi regrouper les observations en lots pour améliorer la normalité.

Expériences pilotes pour estimer β^*?

En général, pour avoir un estimateur sans biais de $\boldsymbol{\beta}^*$, on peut faire une expérience pilote de n_0 observations, calculer un estimateur $\hat{\boldsymbol{\beta}}_0$ et l'utiliser pour les $n - n_0$ observations restantes. On obtient:

$$\bar{X}_{\mathrm{cp},n} = \frac{1}{n-n_0} \sum_{i=n_0+1}^{n} (X_i - \hat{\boldsymbol{\beta}}_0^{\mathsf{t}}(\mathbf{C}_i - \boldsymbol{\nu})) \quad \text{et}$$

$$S^2_{\mathrm{cp},n} = \frac{1}{(n-n_0-1)} \sum_{i=n_0+1}^{n} (X_i - \hat{\boldsymbol{\beta}}_0^{\mathsf{t}}(\mathbf{C}_i - \boldsymbol{\nu}) - \bar{X}_{\mathrm{cp},n})^2.$$

Expériences pilotes pour estimer β^*?

En général, pour avoir un estimateur sans biais de $\boldsymbol{\beta}^*$, on peut faire une expérience pilote de n_0 observations, calculer un estimateur $\hat{\boldsymbol{\beta}}_0$ et l'utiliser pour les $n - n_0$ observations restantes. On obtient:

$$\bar{X}_{\mathrm{cp},n} = \frac{1}{n - n_0} \sum_{i=n_0+1}^{n} (X_i - \hat{\boldsymbol{\beta}}_0^{\mathsf{t}} (\mathbf{C}_i - \boldsymbol{\nu})) \quad \text{et}$$

$$S^2_{\mathrm{cp},n} = \frac{1}{(n - n_0 - 1)} \sum_{i=n_0+1}^{n} (X_i - \hat{\boldsymbol{\beta}}_0^{\mathsf{t}} (\mathbf{C}_i - \boldsymbol{\nu}) - \bar{X}_{\mathrm{cp},n})^2.$$

On a $\mathbb{E}[\bar{X}_{\mathrm{cp},n}] = \mu$ et $\mathbb{E}[S^2_{\mathrm{cp},n}/(n - n_0)] = \mathrm{Var}[\bar{X}_{\mathrm{cp},n}]$.

Expériences pilotes pour estimer β^*?

En général, pour avoir un estimateur sans biais de $\boldsymbol{\beta}^*$, on peut faire une expérience pilote de n_0 observations, calculer un estimateur $\hat{\boldsymbol{\beta}}_0$ et l'utiliser pour les $n - n_0$ observations restantes. On obtient:

$$\bar{X}_{\mathrm{cp},n} = \frac{1}{n-n_0} \sum_{i=n_0+1}^{n} (X_i - \hat{\boldsymbol{\beta}}_0^{\mathsf{t}}(\mathbf{C}_i - \boldsymbol{\nu})) \quad \text{et}$$

$$S^2_{\mathrm{cp},n} = \frac{1}{(n-n_0-1)} \sum_{i=n_0+1}^{n} (X_i - \hat{\boldsymbol{\beta}}_0^{\mathsf{t}}(\mathbf{C}_i - \boldsymbol{\nu}) - \bar{X}_{\mathrm{cp},n})^2.$$

On a $\mathbb{E}[\bar{X}_{\mathrm{cp},n}] = \mu$ et $\mathbb{E}[S^2_{\mathrm{cp},n}/(n-n_0)] = \mathrm{Var}[\bar{X}_{\mathrm{cp},n}]$.

Mais sous l'hypothèse de normalité,

$$\frac{\mathrm{Var}[\bar{X}_{\mathrm{cp},n}]}{\mathrm{Var}[\bar{X}_{\mathrm{ce},n}]} = \frac{n(n-q-2)(n_0-2)}{(n-n_0)(n-2)(n_0-q-2)} > 1.$$

C'est donc inefficace.

Variance des VCs connue

Si on connait $\mathbf{\Sigma}_{\mathrm{C}}$, faut-il remplacer $\hat{\boldsymbol{\beta}}_n$ par

$$\ddot{\boldsymbol{\beta}}_n = \hat{\mathbf{\Sigma}}_{\mathrm{CX}} \mathbf{\Sigma}_{\mathrm{C}}^{-1}?$$

Variance des VCs connue

Si on connait $\mathbf{\Sigma}_{\mathrm{C}}$, faut-il remplacer $\hat{\boldsymbol{\beta}}_n$ par

$$\ddot{\boldsymbol{\beta}}_n = \hat{\mathbf{\Sigma}}_{\mathrm{CX}} \mathbf{\Sigma}_{\mathrm{C}}^{-1} ?$$

Dans le cas normal, c'est moins bon!

Variance des VCs connue

Si on connait $\mathbf{\Sigma}_{\mathrm{C}}$, faut-il remplacer $\hat{\boldsymbol{\beta}}_n$ par

$$\ddot{\boldsymbol{\beta}}_n = \hat{\mathbf{\Sigma}}_{\mathrm{CX}} \mathbf{\Sigma}_{\mathrm{C}}^{-1} ?$$

Dans le cas normal, c'est moins bon!

$\hat{\mathbf{\Sigma}}_{\mathrm{C}}$ agit comme une VC non linéaire.

Exemple: option asiatique

Revenu net

$$X = e^{-rT} \max\left(0, \frac{1}{d}\sum_{j=1}^{d} S(t_j) - K\right).$$

où S est un mouvement brownien géométrique.

Exemple: option asiatique

Revenu net

$$X = e^{-rT} \max\left(0,\ \frac{1}{d}\sum_{j=1}^{d} S(t_j) - K\right).$$

où S est un mouvement brownien géométrique.
On ne connait pas $\mathbb{E}[X]$ mais on connait $\mathbb{E}[C^{(1)}]$ et $\mathbb{E}[C^{(2)}]$ où

$$C^{(1)} = e^{-rT} \max\left(0,\ \prod_{j=1}^{d} S(t_j)^{1/d} - K\right) \quad \text{et} \quad C^{(2)} = \sum_{j=1}^{d} S(t_j)$$

sont souvent fortement corrélées avec X.

Exemple: option asiatique

Revenu net

$$X = e^{-rT} \max\left(0, \frac{1}{d}\sum_{j=1}^{d} S(t_j) - K\right).$$

où S est un mouvement brownien géométrique.
On ne connait pas $\mathbb{E}[X]$ mais on connait $\mathbb{E}[C^{(1)}]$ et $\mathbb{E}[C^{(2)}]$ où

$$C^{(1)} = e^{-rT} \max\left(0, \prod_{j=1}^{d} S(t_j)^{1/d} - K\right) \quad \text{et} \quad C^{(2)} = \sum_{j=1}^{d} S(t_j)$$

sont souvent fortement corrélées avec X.

On peut donc les utiliser comme VCs.

Exemple: option asiatique

Revenu net

$$X = e^{-rT} \max\left(0, \frac{1}{d}\sum_{j=1}^{d} S(t_j) - K\right).$$

où S est un mouvement brownien géométrique.
On ne connait pas $\mathbb{E}[X]$ mais on connait $\mathbb{E}[C^{(1)}]$ et $\mathbb{E}[C^{(2)}]$ où

$$C^{(1)} = e^{-rT} \max\left(0, \prod_{j=1}^{d} S(t_j)^{1/d} - K\right) \quad \text{et} \quad C^{(2)} = \sum_{j=1}^{d} S(t_j)$$

sont souvent fortement corrélées avec X.

On peut donc les utiliser comme VCs.

À essayer en exercice.

Exemple: Option d'achat "down and out" avec barrière ℓ

Le gain est

$$X = e^{-rT}\max(0,\, S(T) - K)\, \mathbb{I}[\min(S(t_1), \ldots, S(t_d)) > \ell],$$

où $0 \le t_1 < \cdots < t_d \le T$ et ℓ sont fixés.

Exemple: Option d'achat "down and out" avec barrière ℓ

Le gain est

$$X = e^{-rT} \max(0,\ S(T) - K)\, \mathbb{I}[\min(S(t_1), \ldots, S(t_d)) > \ell],$$

où $0 \le t_1 < \cdots < t_d \le T$ et ℓ sont fixés.

Comme VC, on peut utiliser

$$C = e^{-rT} \max(0,\ S(T) - K)\,,$$

le gain pour une option européenne ordinaire.

Exemple: Option d'achat "down and out" avec barrière ℓ

Le gain est

$$X = e^{-rT} \max(0,\, S(T) - K)\, \mathbb{I}[\min(S(t_1), \ldots, S(t_d)) > \ell],$$

où $0 \le t_1 < \cdots < t_d \le T$ et ℓ sont fixés.

Comme VC, on peut utiliser

$$C = e^{-rT} \max(0,\, S(T) - K)\,,$$

le gain pour une option européenne ordinaire.

Sera fortement corrélé surtout si la probabilité de "knock-out" est faible.

Exemple: Option d'achat "down and out" avec barrière ℓ

Le gain est

$$X = e^{-rT} \max(0,\ S(T) - K)\, \mathbb{I}[\min(S(t_1), \ldots, S(t_d)) > \ell],$$

où $0 \leq t_1 < \cdots < t_d \leq T$ et ℓ sont fixés.

Comme VC, on peut utiliser

$$C = e^{-rT} \max(0,\ S(T) - K)\,,$$

le gain pour une option européenne ordinaire.

Sera fortement corrélé surtout si la probabilité de "knock-out" est faible.

Autre possibilité:

$$C = e^{-rT} \max(0,\ S(T) - K)\, \mathbb{I}[\min_{0 \leq t \leq T} S(t) > \ell]$$

(on sait comment calculer son espérance).

VCs non linéaires et fonctions de plusieurs espérances

Cadre général: On veut estimer $g(\boldsymbol{\mu})$ où $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)^{\mathsf{t}}$ par $h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n)$, où $(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) \Rightarrow (\boldsymbol{\mu}, \boldsymbol{\nu})$ quand $n \to \infty$, $h(\mathbf{x}, \boldsymbol{\nu}) = g(\mathbf{x})$, et h est continûment différentiable dans un voisinage de $(\boldsymbol{\mu}, \boldsymbol{\nu})$.

VCs non linéaires et fonctions de plusieurs espérances

Cadre général: On veut estimer $g(\boldsymbol{\mu})$ où $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)^{\mathsf{t}}$ par $h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n)$, où $(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) \Rightarrow (\boldsymbol{\mu}, \boldsymbol{\nu})$ quand $n \to \infty$, $h(\mathbf{x}, \boldsymbol{\nu}) = g(\mathbf{x})$, et h est continûment différentiable dans un voisinage de $(\boldsymbol{\mu}, \boldsymbol{\nu})$.

Hypothèse CV2:

$$\sqrt{n}\begin{pmatrix} \bar{\mathbf{X}}_n - \boldsymbol{\mu} \\ \bar{\mathbf{C}}_n - \boldsymbol{\nu} \end{pmatrix} \Rightarrow N(0, \boldsymbol{\Sigma})$$

où

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathrm{X}} & \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}} \\ \boldsymbol{\Sigma}_{\mathrm{CX}} & \boldsymbol{\Sigma}_{\mathrm{C}} \end{pmatrix}$$

est finie et définie positive, $\boldsymbol{\Sigma}_{\mathrm{X}}$ est $d \times d$, et $\boldsymbol{\Sigma}_{\mathrm{C}}$ est $q \times q$.

VCs non linéaires et fonctions de plusieurs espérances

Cadre général: On veut estimer $g(\boldsymbol{\mu})$ où $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)^{\mathsf{t}}$ par $h(\bar{\mathbf{X}}_n,\, \bar{\mathbf{C}}_n)$, où $(\bar{\mathbf{X}}_n,\, \bar{\mathbf{C}}_n) \Rightarrow (\boldsymbol{\mu}, \boldsymbol{\nu})$ quand $n \to \infty$, $h(\mathbf{x}, \boldsymbol{\nu}) = g(\mathbf{x})$, et h est continûment différentiable dans un voisinage de $(\boldsymbol{\mu}, \boldsymbol{\nu})$.

Hypothèse CV2:

$$\sqrt{n} \begin{pmatrix} \bar{\mathbf{X}}_n - \boldsymbol{\mu} \\ \bar{\mathbf{C}}_n - \boldsymbol{\nu} \end{pmatrix} \Rightarrow N(0,\, \boldsymbol{\Sigma})$$

où

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathrm{X}} & \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}} \\ \boldsymbol{\Sigma}_{\mathrm{CX}} & \boldsymbol{\Sigma}_{\mathrm{C}} \end{pmatrix}$$

est finie et définie positive, $\boldsymbol{\Sigma}_{\mathrm{X}}$ est $d \times d$, et $\boldsymbol{\Sigma}_{\mathrm{C}}$ est $q \times q$.

Théorème. Sous CV2, quand $n \to \infty$, on a

$$\begin{aligned} \sqrt{n}[h(\bar{\mathbf{X}}_n,\, \bar{\mathbf{C}}_n) - g(\boldsymbol{\mu})] &= \sqrt{n}[h(\bar{\mathbf{X}}_n,\, \bar{\mathbf{C}}_n) - h(\boldsymbol{\mu}, \boldsymbol{\nu})] \\ \Rightarrow\quad & \sqrt{n}\nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})' \begin{pmatrix} \bar{\mathbf{X}}_n - \boldsymbol{\mu} \\ \bar{\mathbf{C}}_n - \boldsymbol{\nu} \end{pmatrix} \\ =\quad & \sqrt{n}[\nabla_{\boldsymbol{\mu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu})] \end{aligned}$$

$$
\begin{aligned}
&= \sqrt{n}[\nabla g(\boldsymbol{\mu})^{\mathsf{t}}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu},\boldsymbol{\nu})^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu})] \\
&\Rightarrow N(0, \sigma_{\mathrm{h}}^2)
\end{aligned}
$$

où

$$
\begin{aligned}
\sigma_{\mathrm{h}}^2 &= \nabla h(\boldsymbol{\mu},\boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu},\boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{X}} \nabla g(\boldsymbol{\mu}) + 2\nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu},\boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu},\boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{C}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu},\boldsymbol{\nu}).
\end{aligned}
$$

$$
\begin{aligned}
&= \sqrt{n}[\nabla g(\boldsymbol{\mu})^{\mathsf{t}}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu})] \\
&\Rightarrow N(0, \sigma_{\mathrm{h}}^2)
\end{aligned}
$$

où

$$
\begin{aligned}
\sigma_{\mathrm{h}}^2 &= \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{X}} \nabla g(\boldsymbol{\mu}) + 2 \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{C}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned}
$$

Preuve: développement en série de Taylor de $h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n)$ autour de $h(\boldsymbol{\mu}, \boldsymbol{\nu}) = g(\boldsymbol{\mu})$, ou encore utiliser le théorème delta.

$$
\begin{aligned}
&= \sqrt{n}[\nabla g(\boldsymbol{\mu})^{\mathsf{t}}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu})] \\
&\Rightarrow N(0, \sigma_{\mathrm{h}}^2)
\end{aligned}
$$

où

$$
\begin{aligned}
\sigma_{\mathrm{h}}^2 &= \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{X}} \nabla g(\boldsymbol{\mu}) + 2 \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{C}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned}
$$

Preuve: développement en série de Taylor de $h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n)$ autour de $h(\boldsymbol{\mu}, \boldsymbol{\nu}) = g(\boldsymbol{\mu})$, ou encore utiliser le théorème delta.

Sans les VCs, σ_{h}^2 devient égal à $\sigma_{\mathrm{g}}^2 = \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{X}} \nabla g(\boldsymbol{\mu})$.

$$
\begin{aligned}
&= \sqrt{n}[\nabla g(\boldsymbol{\mu})^{\mathsf{t}}(\bar{\mathbf{X}}_n - \boldsymbol{\mu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu})] \\
&\Rightarrow N(0, \sigma_{\mathrm{h}}^2)
\end{aligned}
$$

où

$$
\begin{aligned}
\sigma_{\mathrm{h}}^2 &= \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{X}} \nabla g(\boldsymbol{\mu}) + 2\nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{C}} \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned}
$$

Preuve: développement en série de Taylor de $h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n)$ autour de $h(\boldsymbol{\mu}, \boldsymbol{\nu}) = g(\boldsymbol{\mu})$, ou encore utiliser le théorème delta.

Sans les VCs, σ_{h}^2 devient égal à $\sigma_{\mathrm{g}}^2 = \nabla g(\boldsymbol{\mu})^{\mathsf{t}} \boldsymbol{\Sigma}_{\mathrm{X}} \nabla g(\boldsymbol{\mu})$.
La variance est réduite ssi $\sigma_{\mathrm{h}}^2 < \sigma_{\mathrm{g}}^2$.

Cas particulier des VCs linéaires:

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) = g(\bar{\mathbf{X}}_n) - \boldsymbol{\beta}^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

Cas particulier des VCs linéaires:

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) = g(\bar{\mathbf{X}}_n) - \boldsymbol{\beta}^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

La variance est minimisée en prenant

$$\boldsymbol{\beta}^* = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}\nabla g(\boldsymbol{\mu})$$

et la variance minimale est

$$\sigma_{\mathrm{c}}^2 = \nabla g(\boldsymbol{\mu})^{\mathsf{t}}(\boldsymbol{\Sigma}_{\mathrm{X}} - \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}}\boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}})\nabla g(\boldsymbol{\mu}) = (1 - R_{\mathrm{CX}}^2)\sigma_{\mathrm{g}}^2.$$

Cas particulier des VCs linéaires:

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) = g(\bar{\mathbf{X}}_n) - \boldsymbol{\beta}^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

La variance est minimisée en prenant

$$\boldsymbol{\beta}^* = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}\nabla g(\boldsymbol{\mu})$$

et la variance minimale est

$$\sigma_{\mathrm{c}}^2 = \nabla g(\boldsymbol{\mu})^{\mathsf{t}}(\boldsymbol{\Sigma}_{\mathrm{X}} - \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathsf{t}}\boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}})\nabla g(\boldsymbol{\mu}) = (1 - R_{\mathrm{CX}}^2)\sigma_{\mathrm{g}}^2.$$

Dans le cas général, on a

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) \quad = \quad g(\bar{\mathbf{X}}_n) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathsf{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}) + o_p(n^{-1/2}).$$

Cas particulier des VCs linéaires:

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) = g(\bar{\mathbf{X}}_n) - \boldsymbol{\beta}^{\mathrm{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

La variance est minimisée en prenant

$$\boldsymbol{\beta}^* = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}\nabla g(\boldsymbol{\mu})$$

et la variance minimale est

$$\sigma_{\mathrm{c}}^2 = \nabla g(\boldsymbol{\mu})^{\mathrm{t}}(\boldsymbol{\Sigma}_{\mathrm{X}} - \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathrm{t}}\boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}})\nabla g(\boldsymbol{\mu}) = (1 - R_{\mathrm{CX}}^2)\sigma_{\mathrm{g}}^2.$$

Dans le cas général, on a

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) \quad = \quad g(\bar{\mathbf{X}}_n) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathrm{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}) + o_p(n^{-1/2}).$$

L'estimateur avec VCs non linéaires est donc asymptotiquement équivalent à utiliser $\bar{\mathbf{C}}_n$ comme VC linéaire, avec $\boldsymbol{\beta} = -\nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})$.

Cas particulier des VCs linéaires:

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) = g(\bar{\mathbf{X}}_n) - \boldsymbol{\beta}^{\mathrm{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

La variance est minimisée en prenant

$$\boldsymbol{\beta}^* = \boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}}\nabla g(\boldsymbol{\mu})$$

et la variance minimale est

$$\sigma_{\mathrm{c}}^2 = \nabla g(\boldsymbol{\mu})^{\mathrm{t}}(\boldsymbol{\Sigma}_{\mathrm{X}} - \boldsymbol{\Sigma}_{\mathrm{CX}}^{\mathrm{t}}\boldsymbol{\Sigma}_{\mathrm{C}}^{-1}\boldsymbol{\Sigma}_{\mathrm{CX}})\nabla g(\boldsymbol{\mu}) = (1 - R_{\mathrm{CX}}^2)\sigma_{\mathrm{g}}^2.$$

Dans le cas général, on a

$$h(\bar{\mathbf{X}}_n, \bar{\mathbf{C}}_n) \quad = \quad g(\bar{\mathbf{X}}_n) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^{\mathrm{t}}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}) + o_p(n^{-1/2}).$$

L'estimateur avec VCs non linéaires est donc asymptotiquement équivalent à utiliser $\bar{\mathbf{C}}_n$ comme VC linéaire, avec $\boldsymbol{\beta} = -\nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})$.

Pas optimal pour $n \to \infty$, mais peut être utile pour n fini.

Exemple: estimation d'un quotient d'espérances.
$g(\boldsymbol{\mu}) = \mu_1/\mu_2$ où $\boldsymbol{\mu} = (\mu_1, \mu_2)^{\mathrm{t}}$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^{\mathrm{t}}$.

Exemple: estimation d'un quotient d'espérances.
$g(\boldsymbol{\mu}) = \mu_1/\mu_2$ où $\boldsymbol{\mu} = (\mu_1, \mu_2)^{\mathrm{t}}$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^{\mathrm{t}}$.
Estimateur de μ_1/μ_2 avec VCs linéaires:

$$\frac{\hat{\mu}_1}{\hat{\mu}_2} - \hat{\boldsymbol{\beta}}_n(\bar{\mathbf{C}}_n - \boldsymbol{\nu}) = \frac{\hat{\mu}_1}{\hat{\mu}_2} - (1/\hat{\mu}_2, -\hat{\mu}_1/\hat{\mu}_2^2)\hat{\boldsymbol{\Sigma}}_{\mathrm{CX}}^{\mathrm{t}}\hat{\boldsymbol{\Sigma}}_{\mathrm{C}}^{-1}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

Exemple: estimation d'un quotient d'espérances.
$g(\boldsymbol{\mu}) = \mu_1/\mu_2$ où $\boldsymbol{\mu} = (\mu_1, \mu_2)^{\mathrm{t}}$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^{\mathrm{t}}$.
Estimateur de μ_1/μ_2 avec VCs linéaires:

$$\frac{\hat{\mu}_1}{\hat{\mu}_2} - \hat{\boldsymbol{\beta}}_n(\bar{\mathbf{C}}_n - \boldsymbol{\nu}) = \frac{\hat{\mu}_1}{\hat{\mu}_2} - (1/\hat{\mu}_2, -\hat{\mu}_1/\hat{\mu}_2^2)\hat{\boldsymbol{\Sigma}}_{\mathrm{CX}}^{\mathrm{t}}\hat{\boldsymbol{\Sigma}}_{\mathrm{C}}^{-1}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

Exemples de VCs non linéaires (en une dimension):

$$\begin{aligned} h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)\bar{C}_n/\nu \\ h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)\nu/\bar{C}_n \\ h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)^{\bar{C}_n/\nu} \\ h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)^{\nu/\bar{C}_n}, \end{aligned}$$

où $\mathbb{E}[C_n] = \nu$. Biaisé, mais réduit le MSE dans certains cas.

Exemple: estimation d'un quotient d'espérances.
$g(\boldsymbol{\mu}) = \mu_1/\mu_2$ où $\boldsymbol{\mu} = (\mu_1, \mu_2)^{\mathrm{t}}$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^{\mathrm{t}}$.
Estimateur de μ_1/μ_2 avec VCs linéaires:

$$\frac{\hat{\mu}_1}{\hat{\mu}_2} - \hat{\boldsymbol{\beta}}_n(\bar{\mathbf{C}}_n - \boldsymbol{\nu}) = \frac{\hat{\mu}_1}{\hat{\mu}_2} - (1/\hat{\mu}_2, -\hat{\mu}_1/\hat{\mu}_2^2)\hat{\boldsymbol{\Sigma}}_{\mathrm{CX}}^{\mathrm{t}}\hat{\boldsymbol{\Sigma}}_{\mathrm{C}}^{-1}(\bar{\mathbf{C}}_n - \boldsymbol{\nu}).$$

Exemples de VCs non linéaires (en une dimension):

$$\begin{aligned}
h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)\bar{C}_n/\nu \\
h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)\nu/\bar{C}_n \\
h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)^{\bar{C}_n/\nu} \\
h(\bar{X}_n, \bar{C}_n) &= g(\bar{X}_n)^{\nu/\bar{C}_n},
\end{aligned}$$

où $\mathbb{E}[C_n] = \nu$. Biaisé, mais réduit le MSE dans certains cas.
Utilisé lorsque l'on remplace $\boldsymbol{\Sigma}_{\mathrm{C}}$ (connu) par $\hat{\boldsymbol{\Sigma}}_{\mathrm{C}}$ pour estimer $\hat{\boldsymbol{\beta}}^*$.

Appariement des moments

Consiste à modifier l'échantillon de v.a. pour que sa moyenne, variance, etc., correspondent exactement aux valeurs théoriques.

Appariement des moments

Consiste à modifier l'échantillon de v.a. pour que sa moyenne, variance, etc., correspondent exactement aux valeurs théoriques.

Souvent utilisé en économie et finance.
Ressemble à l'utilisation des VC non linéaires.
Donne des estimateurs biaisés mais réduit parfois le MSE.

Appariement des moments

Consiste à modifier l'échantillon de v.a. pour que sa moyenne, variance, etc., correspondent exactement aux valeurs théoriques.

Souvent utilisé en économie et finance.
Ressemble à l'utilisation des VC non linéaires.
Donne des estimateurs biaisés mais réduit parfois le MSE.

Supposons que chaque simulation dépend d'une v.a. Z_i.
(Peut se généraliser à un vecteur.)
On connait $\mu_{\mathrm{z}} = \mathbb{E}[Z_i]$, $\sigma_{\mathrm{z}}^2 = \mathrm{Var}[Z_i]$, etc.

Appariement des moments

Consiste à modifier l'échantillon de v.a. pour que sa moyenne, variance, etc., correspondent exactement aux valeurs théoriques.

Souvent utilisé en économie et finance.
Ressemble à l'utilisation des VC non linéaires.
Donne des estimateurs biaisés mais réduit parfois le MSE.

Supposons que chaque simulation dépend d'une v.a. Z_i.
(Peut se généraliser à un vecteur.)
On connait $\mu_{\mathrm{z}} = \mathbb{E}[Z_i]$, $\sigma_{\mathrm{z}}^2 = \mathrm{Var}[Z_i]$, etc.

Pour n répétitions, on a $Z_1, \ldots, Z_n$ i.i.d.

Appariement des moments

Consiste à modifier l'échantillon de v.a. pour que sa moyenne, variance, etc., correspondent exactement aux valeurs théoriques.

Souvent utilisé en économie et finance.
Ressemble à l'utilisation des VC non linéaires.
Donne des estimateurs biaisés mais réduit parfois le MSE.

Supposons que chaque simulation dépend d'une v.a. Z_i.
(Peut se généraliser à un vecteur.)
On connait $\mu_{\mathrm{z}} = \mathbb{E}[Z_i]$, $\sigma_{\mathrm{z}}^2 = \mathrm{Var}[Z_i]$, etc.

Pour n répétitions, on a $Z_1, \ldots, Z_n$ i.i.d.
Pour ajuster seulement la moyenne, remplacer Z_i par

$$\tilde{Z}_i = Z_i + \mu_{\mathrm{z}} - \bar{Z}_n.$$

Appariement des moments

Consiste à modifier l'échantillon de v.a. pour que sa moyenne, variance, etc., correspondent exactement aux valeurs théoriques.

Souvent utilisé en économie et finance.
Ressemble à l'utilisation des VC non linéaires.
Donne des estimateurs biaisés mais réduit parfois le MSE.

Supposons que chaque simulation dépend d'une v.a. Z_i.
(Peut se généraliser à un vecteur.)
On connait $\mu_{\mathrm{z}} = \mathbb{E}[Z_i]$, $\sigma_{\mathrm{z}}^2 = \mathrm{Var}[Z_i]$, etc.

Pour n répétitions, on a $Z_1, \ldots, Z_n$ i.i.d.
Pour ajuster seulement la moyenne, remplacer Z_i par

$$\tilde{Z}_i = Z_i + \mu_{\mathrm{z}} - \bar{Z}_n.$$

Pour ajuster la moyenne et la variance, remplacer Z_i par

$$\tilde{Z}_i = \mu_{\mathrm{z}} + (Z_i - \bar{Z}_n)\sigma_{\mathrm{z}}/S_{\mathrm{z},n}$$

où $S_{\mathrm{z},n}^2$ est la variance empirique des Z_i.

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

On peut montrer que

$$\sqrt{n}(\bar{X}_{\mathrm{mm},n} - \mu) \;\Rightarrow\; \sqrt{n}(\bar{X}_n - \beta_1 C^{(1)} - \beta_2 C^{(2)} - \mu)$$

quand $n \to \infty$, où $C^{(1)} = 1 - \sigma_{\mathrm{z}}/S_{\mathrm{z},n}$, $C^{(2)} = \bar{Z}_n \sigma_{\mathrm{z}}/S_{\mathrm{z},n} - \mu_{\mathrm{z}}$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
et $\beta_2 = \mathbb{E}[h'(Z_i)]$.

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

On peut montrer que

$$\sqrt{n}(\bar{X}_{\mathrm{mm},n} - \mu) \;\Rightarrow\; \sqrt{n}(\bar{X}_n - \beta_1 C^{(1)} - \beta_2 C^{(2)} - \mu)$$

quand $n \to \infty$, où $C^{(1)} = 1 - \sigma_{\mathrm{z}}/S_{\mathrm{z},n}$, $C^{(2)} = \bar{Z}_n \sigma_{\mathrm{z}}/S_{\mathrm{z},n} - \mu_{\mathrm{z}}$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$, et $\beta_2 = \mathbb{E}[h'(Z_i)]$.

Asymptotiquement équivalent à utiliser $C^{(1)}$ et $C^{(2)}$ comme VCs linéaires avec coefficients β_1 et β_2.

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n} \sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

On peut montrer que

$$\sqrt{n}(\bar{X}_{\mathrm{mm},n} - \mu) \;\Rightarrow\; \sqrt{n}(\bar{X}_n - \beta_1 C^{(1)} - \beta_2 C^{(2)} - \mu)$$

quand $n \to \infty$, où $C^{(1)} = 1 - \sigma_{\mathrm{z}}/S_{\mathrm{z},n}$, $C^{(2)} = \bar{Z}_n \sigma_{\mathrm{z}}/S_{\mathrm{z},n} - \mu_{\mathrm{z}}$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$, et $\beta_2 = \mathbb{E}[h'(Z_i)]$.

Asymptotiquement équivalent à utiliser $C^{(1)}$ et $C^{(2)}$ comme VCs linéaires avec coefficients β_1 et β_2. Pas nécessairement optimal!

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

On peut montrer que

$$\sqrt{n}(\bar{X}_{\mathrm{mm},n} - \mu) \Rightarrow \sqrt{n}(\bar{X}_n - \beta_1 C^{(1)} - \beta_2 C^{(2)} - \mu)$$

quand $n \to \infty$, où $C^{(1)} = 1 - \sigma_{\mathrm{z}}/S_{\mathrm{z},n}$, $C^{(2)} = \bar{Z}_n \sigma_{\mathrm{z}}/S_{\mathrm{z},n} - \mu_{\mathrm{z}}$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$, et $\beta_2 = \mathbb{E}[h'(Z_i)]$.

Asymptotiquement équivalent à utiliser $C^{(1)}$ et $C^{(2)}$ comme VCs linéaires avec coefficients β_1 et β_2. Pas nécessairement optimal!

Des plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(1)}] \neq 0$ en général.

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

On peut montrer que

$$\sqrt{n}(\bar{X}_{\mathrm{mm},n} - \mu) \;\Rightarrow\; \sqrt{n}(\bar{X}_n - \beta_1 C^{(1)} - \beta_2 C^{(2)} - \mu)$$

quand $n \to \infty$, où $C^{(1)} = 1 - \sigma_{\mathrm{z}}/S_{\mathrm{z},n}$, $C^{(2)} = \bar{Z}_n\sigma_{\mathrm{z}}/S_{\mathrm{z},n} - \mu_{\mathrm{z}}$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$, et $\beta_2 = \mathbb{E}[h'(Z_i)]$.

Asymptotiquement équivalent à utiliser $C^{(1)}$ et $C^{(2)}$ comme VCs linéaires avec coefficients β_1 et β_2. Pas nécessairement optimal!

Des plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(1)}] \neq 0$ en général.

On pourrait utiliser les VCs $\bar{Z}_n - \mu_{\mathrm{z}}$ et $S^2_{\mathrm{z},n} - \sigma^2_{\mathrm{z}}$ à la place.

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

On peut montrer que

$$\sqrt{n}(\bar{X}_{\mathrm{mm},n} - \mu) \;\Rightarrow\; \sqrt{n}(\bar{X}_n - \beta_1 C^{(1)} - \beta_2 C^{(2)} - \mu)$$

quand $n \to \infty$, où $C^{(1)} = 1 - \sigma_{\mathrm{z}}/S_{\mathrm{z},n}$, $C^{(2)} = \bar{Z}_n\sigma_{\mathrm{z}}/S_{\mathrm{z},n} - \mu_{\mathrm{z}}$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$, et $\beta_2 = \mathbb{E}[h'(Z_i)]$.

Asymptotiquement équivalent à utiliser $C^{(1)}$ et $C^{(2)}$ comme VCs linéaires avec coefficients β_1 et β_2. Pas nécessairement optimal!

Des plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(1)}] \neq 0$ en général.

On pourrait utiliser les VCs $\bar{Z}_n - \mu_{\mathrm{z}}$ et $S^2_{\mathrm{z},n} - \sigma^2_{\mathrm{z}}$ à la place.

Autre difficulté: comment estimer $\mathrm{Var}[\bar{X}_{\mathrm{mm},n}]$?

Supposons que l'on veut estimer $\mu = \mathbb{E}[h(Z_i)]$, où h'' est continue.
On pose $X_i = h(Z_i)$, $X_{\mathrm{mm},i} = h(\tilde{Z}_i)$,

$$\bar{X}_{\mathrm{mm},n} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^{n} X_{\mathrm{mm},i}.$$

On peut montrer que

$$\sqrt{n}(\bar{X}_{\mathrm{mm},n} - \mu) \;\Rightarrow\; \sqrt{n}(\bar{X}_n - \beta_1 C^{(1)} - \beta_2 C^{(2)} - \mu)$$

quand $n \to \infty$, où $C^{(1)} = 1 - \sigma_{\mathrm{z}}/S_{\mathrm{z},n}$, $C^{(2)} = \bar{Z}_n \sigma_{\mathrm{z}}/S_{\mathrm{z},n} - \mu_{\mathrm{z}}$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$, et $\beta_2 = \mathbb{E}[h'(Z_i)]$.

Asymptotiquement équivalent à utiliser $C^{(1)}$ et $C^{(2)}$ comme VCs linéaires avec coefficients β_1 et β_2. Pas nécessairement optimal!

Des plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(1)}] \neq 0$ en général.

On pourrait utiliser les VCs $\bar{Z}_n - \mu_{\mathrm{z}}$ et $S^2_{\mathrm{z},n} - \sigma^2_{\mathrm{z}}$ à la place.

Autre difficulté: comment estimer $\mathrm{Var}[\bar{X}_{\mathrm{mm},n}]$?
On peut générer m copies indépendantes de cet estimateur.