

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 $\mathbf{C} = (C^{(1)}, \dots, C^{(q)})^t$ des VCs corrélées avec X , d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

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 $\mathbf{C} = (C^{(1)}, \dots, C^{(q)})^t$ des VCs corrélées avec X , d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

L'estimateur avec VC est:

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

où $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^t$ (des constantes).

On a $\mathbb{E}[X_c] = \mathbb{E}[X] = \mu$.

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 $\mathbf{C} = (C^{(1)}, \dots, C^{(q)})^t$ des VCs corrélées avec X , d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

L'estimateur avec VC est:

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

où $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^t$ (des constantes).

On a $\mathbb{E}[X_c] = \mathbb{E}[X] = \mu$.

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 $\mathbf{C} = (C^{(1)}, \dots, C^{(q)})^t$ des VCs corrélées avec X , d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

L'estimateur avec VC est:

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

où $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^t$ (des constantes).

On a $\mathbb{E}[X_c] = \mathbb{E}[X] = \mu$.

Comment choisir $\boldsymbol{\beta}$?

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

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 $\mathbf{C} = (C^{(1)}, \dots, C^{(q)})^t$ des VCs corrélées avec X , d'espérance connue $\mathbb{E}[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

L'estimateur avec VC est:

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

où $\boldsymbol{\beta} = (\beta_1, \dots, \beta_q)^t$ (des constantes).

On a $\mathbb{E}[X_c] = \mathbb{E}[X] = \mu$.

Comment choisir $\boldsymbol{\beta}$?

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

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

On a alors

$$\text{Var}[X_c] = \text{Var}[X] + \boldsymbol{\beta}^t \boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\beta}^t \boldsymbol{\Sigma}_{CX}.$$

On a alors

$$\text{Var}[X_c] = \text{Var}[X] + \boldsymbol{\beta}^t \boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\beta}^t \boldsymbol{\Sigma}_{CX}.$$

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

$$0 = \nabla_{\boldsymbol{\beta}} \text{Var}[X_c] = 2\boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\Sigma}_{CX}.$$

On a alors

$$\text{Var}[X_c] = \text{Var}[X] + \boldsymbol{\beta}^t \boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\beta}^t \boldsymbol{\Sigma}_{CX}.$$

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

$$0 = \nabla_{\boldsymbol{\beta}} \text{Var}[X_c] = 2\boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\Sigma}_{CX}.$$

Le minimum est donc atteint pour

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

On a alors

$$\text{Var}[X_c] = \text{Var}[X] + \boldsymbol{\beta}^t \boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\beta}^t \boldsymbol{\Sigma}_{CX}.$$

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

$$0 = \nabla_{\boldsymbol{\beta}} \text{Var}[X_c] = 2\boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\Sigma}_{CX}.$$

Le minimum est donc atteint pour

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

qui donne la variance minimale

$$\text{Var}[X_c] = (1 - R_{CX}^2) \text{Var}[X] \stackrel{\text{def}}{=} \sigma_c^2,$$

où

$$R_{CX}^2 = \frac{\boldsymbol{\Sigma}_{CX}^t \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{CX}}{\text{Var}[X]}$$

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

On a alors

$$\text{Var}[X_c] = \text{Var}[X] + \boldsymbol{\beta}^t \boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\beta}^t \boldsymbol{\Sigma}_{CX}.$$

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

$$0 = \nabla_{\boldsymbol{\beta}} \text{Var}[X_c] = 2\boldsymbol{\Sigma}_C \boldsymbol{\beta} - 2\boldsymbol{\Sigma}_{CX}.$$

Le minimum est donc atteint pour

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

qui donne la variance minimale

$$\text{Var}[X_c] = (1 - R_{CX}^2) \text{Var}[X] \stackrel{\text{def}}{=} \sigma_c^2,$$

où

$$R_{CX}^2 = \frac{\boldsymbol{\Sigma}_{CX}^t \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{CX}}{\text{Var}[X]}$$

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

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 $X^{(0)}, \dots, X^{(q)}$ des estimateurs sans biais de μ . Posons

$$X_c = \sum_{\ell=0}^q \beta_{\ell} X^{(\ell)} = X^{(0)} - \sum_{\ell=1}^q \beta_{\ell} (X^{(0)} - X^{(\ell)})$$

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 $X^{(0)}, \dots, X^{(q)}$ des estimateurs sans biais de μ . Posons

$$X_c = \sum_{\ell=0}^q \beta_{\ell} X^{(\ell)} = X^{(0)} - \sum_{\ell=1}^q \beta_{\ell} (X^{(0)} - X^{(\ell)})$$

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

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

Estimation de β^* : propriétés asymptotiques

En pratique, on ne connaît pas $\beta^* = \Sigma_C^{-1} \Sigma_{CX}$
(parfois Σ_C , mais jamais Σ_{CX}).

Estimation de β^* : propriétés asymptotiques

En pratique, on ne connaît pas $\beta^* = \Sigma_C^{-1} \Sigma_{CX}$
(parfois Σ_C , mais jamais Σ_{CX}).

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

Estimation de β^* : propriétés asymptotiques

En pratique, on ne connaît pas $\beta^* = \Sigma_C^{-1} \Sigma_{CX}$
(parfois Σ_C , mais jamais Σ_{CX}).

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

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

et

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

Estimation de β^* : propriétés asymptotiques

En pratique, on ne connaît pas $\beta^* = \Sigma_C^{-1} \Sigma_{CX}$ (parfois Σ_C , mais jamais Σ_{CX}).

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

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

et

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

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

$$\sqrt{n}(\bar{X}_{c,n} - \bar{X}_{ce,n}) \Rightarrow 0,$$

$$S_{ce,n}^2 \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^n (X_{ce,i} - \bar{X}_{ce,n})^2 \Rightarrow \sigma_c^2,$$

$$\frac{\sqrt{n}(\bar{X}_{ce,n} - \mu)}{S_{ce,n}} \Rightarrow \frac{\sqrt{n}(\bar{X}_{c,n} - \mu)}{\sigma_c} \Rightarrow N(0, 1).$$

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

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

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

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

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

$$\hat{\beta}_n = \hat{\Sigma}_C^{-1} \hat{\Sigma}_{CX}$$

où les éléments de $\hat{\Sigma}_C$ et $\hat{\Sigma}_{CX}$ sont

$$\hat{\sigma}_C^{(\ell,k)} = \frac{1}{n-1} \sum_{i=1}^n (C_i^{(\ell)} - \bar{C}_n^{(\ell)})(C_i^{(k)} - \bar{C}_n^{(k)}),$$

$$\hat{\sigma}_{CX}^{(\ell)} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)(C_i^{(\ell)} - \bar{C}_n^{(\ell)}).$$

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

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

$$\hat{\beta}_n = \hat{\Sigma}_C^{-1} \hat{\Sigma}_{CX}$$

où les éléments de $\hat{\Sigma}_C$ et $\hat{\Sigma}_{CX}$ sont

$$\hat{\sigma}_C^{(\ell,k)} = \frac{1}{n-1} \sum_{i=1}^n (C_i^{(\ell)} - \bar{C}_n^{(\ell)})(C_i^{(k)} - \bar{C}_n^{(k)}),$$

$$\hat{\sigma}_{CX}^{(\ell)} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)(C_i^{(\ell)} - \bar{C}_n^{(\ell)}).$$

Variante: remplacer $\bar{C}_n^{(\ell)}$ par $\nu^{(\ell)}$, ou encore $\hat{\sigma}_C^{(\ell,k)}$ par $\sigma_C^{(\ell,k)}$, si connu.

Cas multinormal

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

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

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

Cas multinormal

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

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

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

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

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

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

$$\mathbb{E}[\bar{X}_{\text{ce},n}] = \mu \quad (\text{aucun biais}),$$

$$\mathbb{E}[\tilde{S}_{\text{ce},n}^2/n] = \text{Var}[\bar{X}_{\text{ce},n}] = \frac{n-2}{n-q-2}(1-R_{\text{CX}}^2)\text{Var}[\bar{X}_n],$$

$$\sqrt{n}(\bar{X}_{\text{ce},n} - \mu)/\tilde{S}_{\text{ce},n} \sim \text{Student-t}(n-q-1) \quad (\text{loi exacte}).$$

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

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

$$\mathbb{E}[\bar{X}_{\text{ce},n}] = \mu \quad (\text{aucun biais}),$$

$$\mathbb{E}[\tilde{S}_{\text{ce},n}^2/n] = \text{Var}[\bar{X}_{\text{ce},n}] = \frac{n-2}{n-q-2}(1-R_{\text{CX}}^2)\text{Var}[\bar{X}_n],$$

$$\sqrt{n}(\bar{X}_{\text{ce},n} - \mu)/\tilde{S}_{\text{ce},n} \sim \text{Student-t}(n-q-1) \quad (\text{loi exacte}).$$

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 β^* .

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

$$\mathbb{E}[\bar{X}_{\text{ce},n}] = \mu \quad (\text{aucun biais}),$$

$$\mathbb{E}[\tilde{S}_{\text{ce},n}^2/n] = \text{Var}[\bar{X}_{\text{ce},n}] = \frac{n-2}{n-q-2}(1-R_{\text{CX}}^2)\text{Var}[\bar{X}_n],$$

$$\sqrt{n}(\bar{X}_{\text{ce},n} - \mu)/\tilde{S}_{\text{ce},n} \sim \text{Student-t}(n-q-1) \quad (\text{loi exacte}).$$

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 β^* .

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

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

$$\mathbb{E}[\bar{X}_{\text{ce},n}] = \mu \quad (\text{aucun biais}),$$

$$\mathbb{E}[\tilde{S}_{\text{ce},n}^2/n] = \text{Var}[\bar{X}_{\text{ce},n}] = \frac{n-2}{n-q-2}(1-R_{\text{CX}}^2)\text{Var}[\bar{X}_n],$$

$$\sqrt{n}(\bar{X}_{\text{ce},n} - \mu)/\tilde{S}_{\text{ce},n} \sim \text{Student-t}(n-q-1) \quad (\text{loi exacte}).$$

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 β^* .

Si on a déjà q VC, l'ajout d'une nouvelle VC n'est rentable que si la valeur de $(1 - R_{\text{CX}}^2)$ 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 $(X_i, \mathbf{C}_i^t)^t$ sont i.i.d normaux, alors

$$\mathbb{E}[\bar{X}_{\text{ce},n}] = \mu \quad (\text{aucun biais}),$$

$$\mathbb{E}[\tilde{S}_{\text{ce},n}^2/n] = \text{Var}[\bar{X}_{\text{ce},n}] = \frac{n-2}{n-q-2}(1-R_{\text{CX}}^2)\text{Var}[\bar{X}_n],$$

$$\sqrt{n}(\bar{X}_{\text{ce},n} - \mu)/\tilde{S}_{\text{ce},n} \sim \text{Student-t}(n-q-1) \quad (\text{loi exacte}).$$

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 β^* .

Si on a déjà q VC, l'ajout d'une nouvelle VC n'est rentable que si la valeur de $(1 - R_{\text{CX}}^2)$ 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 β^* , on peut faire une expérience pilote de n_0 observations, calculer un estimateur $\hat{\beta}_0$ et l'utiliser pour les $n - n_0$ observations restantes. On obtient:

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

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

Expériences pilotes pour estimer β^* ?

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

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

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

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

Expériences pilotes pour estimer β^* ?

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

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

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

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

Mais sous l'hypothèse de normalité,

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

C'est donc inefficace.

Variance des VCs connue

Si on connaît Σ_C , faut-il remplacer $\hat{\beta}_n$ par

$$\ddot{\beta}_n = \hat{\Sigma}_{CX} \Sigma_C^{-1}?$$

Variance des VCs connue

Si on connaît Σ_C , faut-il remplacer $\hat{\beta}_n$ par

$$\ddot{\beta}_n = \hat{\Sigma}_{CX} \Sigma_C^{-1}?$$

Dans le cas normal, c'est moins bon!

Variance des VCs connue

Si on connaît Σ_C , faut-il remplacer $\hat{\beta}_n$ par

$$\ddot{\beta}_n = \hat{\Sigma}_{CX} \Sigma_C^{-1}?$$

Dans le cas normal, c'est moins bon!

$\hat{\Sigma}_C$ 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 connaît pas $\mathbb{E}[X]$ mais on connaît $\mathbb{E}[C^{(1)}]$ 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 connaît pas $\mathbb{E}[X]$ mais on connaît $\mathbb{E}[C^{(1)}]$ 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 connaît pas $\mathbb{E}[X]$ mais on connaît $\mathbb{E}[C^{(1)}]$ 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

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

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

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

Le gain est

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

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

Comme VC, on peut utiliser

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

le gain pour une option européenne ordinaire.

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

Le gain est

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

où $0 \leq t_1 < \dots < t_d \leq T$ 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.

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

Le gain est

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

où $0 \leq t_1 < \dots < t_d \leq T$ 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 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^\top$ 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 $n \rightarrow \infty$, $h(\mathbf{x}, \boldsymbol{\nu}) = g(\mathbf{x})$, 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 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^t$ 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 $n \rightarrow \infty$, $h(\mathbf{x}, \boldsymbol{\nu}) = g(\mathbf{x})$, 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

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_X & \boldsymbol{\Sigma}_{CX}^t \\ \boldsymbol{\Sigma}_{CX} & \boldsymbol{\Sigma}_C \end{pmatrix}$$

est finie et définie positive, $\boldsymbol{\Sigma}_X$ est $d \times d$, et $\boldsymbol{\Sigma}_C$ est $q \times q$.

VCs non linéaires et fonctions de plusieurs espérances

Cadre général: On veut estimer $g(\boldsymbol{\mu})$ où $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)^\text{t}$ 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 $n \rightarrow \infty$, $h(\mathbf{x}, \boldsymbol{\nu}) = g(\mathbf{x})$, 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

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_X & \boldsymbol{\Sigma}_{CX}^\text{t} \\ \boldsymbol{\Sigma}_{CX} & \boldsymbol{\Sigma}_C \end{pmatrix}$$

est finie et définie positive, $\boldsymbol{\Sigma}_X$ est $d \times d$, et $\boldsymbol{\Sigma}_C$ est $q \times q$.

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

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

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

où

$$\begin{aligned}
\sigma_h^2 &= \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})^\dagger \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^\dagger \boldsymbol{\Sigma}_X \nabla g(\boldsymbol{\mu}) + 2 \nabla g(\boldsymbol{\mu})^\dagger \boldsymbol{\Sigma}_{CX}^\dagger \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^\dagger \boldsymbol{\Sigma}_C \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned}$$

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

où

$$\begin{aligned}
\sigma_h^2 &= \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})^\top \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_X \nabla g(\boldsymbol{\mu}) + 2 \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_{CX}^\top \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^\top \boldsymbol{\Sigma}_C \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned}$$

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.

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

où

$$\begin{aligned}
\sigma_h^2 &= \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})^\top \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_X \nabla g(\boldsymbol{\mu}) + 2 \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_{CX}^\top \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^\top \boldsymbol{\Sigma}_C \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned}$$

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 $\sigma_g^2 = \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_X \nabla g(\boldsymbol{\mu})$.

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

où

$$\begin{aligned}
\sigma_h^2 &= \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu})^\top \boldsymbol{\Sigma} \nabla h(\boldsymbol{\mu}, \boldsymbol{\nu}) \\
&= \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_X \nabla g(\boldsymbol{\mu}) + 2 \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_{CX}^\top \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}) + \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu})^\top \boldsymbol{\Sigma}_C \nabla_{\boldsymbol{\nu}} h(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned}$$

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 $\sigma_g^2 = \nabla g(\boldsymbol{\mu})^\top \boldsymbol{\Sigma}_X \nabla g(\boldsymbol{\mu})$.

La variance est réduite ssi $\sigma_h^2 < \sigma_g^2$.

Cas particulier des **V**Cs linéaires:

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

Cas particulier des **VCS linéaires**:

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

La variance est minimisée en prenant

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

et la variance minimale est

$$\sigma_c^2 = \nabla g(\boldsymbol{\mu})^\dagger (\boldsymbol{\Sigma}_X - \boldsymbol{\Sigma}_{CX}^\dagger \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{CX}) \nabla g(\boldsymbol{\mu}) = (1 - R_{CX}^2) \sigma_g^2.$$

Cas particulier des **VCS linéaires**:

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

La variance est minimisée en prenant

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

et la variance minimale est

$$\sigma_c^2 = \nabla g(\boldsymbol{\mu})^\dagger (\boldsymbol{\Sigma}_X - \boldsymbol{\Sigma}_{CX}^\dagger \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{CX}) \nabla g(\boldsymbol{\mu}) = (1 - R_{CX}^2) \sigma_g^2.$$

Dans le cas général, on a

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

Cas particulier des **VCS linéaires**:

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

La variance est minimisée en prenant

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

et la variance minimale est

$$\sigma_c^2 = \nabla g(\boldsymbol{\mu})^\dagger (\boldsymbol{\Sigma}_X - \boldsymbol{\Sigma}_{CX}^\dagger \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{CX}) \nabla g(\boldsymbol{\mu}) = (1 - R_{CX}^2) \sigma_g^2.$$

Dans le cas général, on a

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

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**:

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

La variance est minimisée en prenant

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

et la variance minimale est

$$\sigma_c^2 = \nabla g(\boldsymbol{\mu})^\dagger (\boldsymbol{\Sigma}_X - \boldsymbol{\Sigma}_{CX}^\dagger \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{CX}) \nabla g(\boldsymbol{\mu}) = (1 - R_{CX}^2) \sigma_g^2.$$

Dans le cas général, on a

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

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 $n \rightarrow \infty$, mais peut être utile pour n fini.

Exemple: estimation d'un **quotient** d'espérances.

$g(\boldsymbol{\mu}) = \mu_1/\mu_2$ où $\boldsymbol{\mu} = (\mu_1, \mu_2)^t$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^t$.

Exemple: estimation d'un **quotient** d'espérances.

$g(\boldsymbol{\mu}) = \mu_1/\mu_2$ où $\boldsymbol{\mu} = (\mu_1, \mu_2)^t$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^t$.

Estimateur de μ_1/μ_2 avec VCs linéaires:

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

Exemple: estimation d'un **quotient** d'espérances.

$g(\boldsymbol{\mu}) = \mu_1/\mu_2$ où $\boldsymbol{\mu} = (\mu_1, \mu_2)^t$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^t$.

Estimateur de μ_1/μ_2 avec VCs linéaires:

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

Exemples de VCs non linéaires (en une dimension):

$$h(\bar{X}_n, \bar{C}_n) = g(\bar{X}_n) \bar{C}_n / \nu$$

$$h(\bar{X}_n, \bar{C}_n) = g(\bar{X}_n) \nu / \bar{C}_n$$

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

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

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 $\boldsymbol{\mu} = (\mu_1, \mu_2)^t$, et $\nabla g(\boldsymbol{\mu}) = (1/\mu_2, -\mu_1/\mu_2^2)^t$.

Estimateur de μ_1/μ_2 avec VCs linéaires:

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

Exemples de VCs non linéaires (en une dimension):

$$h(\bar{X}_n, \bar{C}_n) = g(\bar{X}_n) \bar{C}_n / \nu$$

$$h(\bar{X}_n, \bar{C}_n) = g(\bar{X}_n) \nu / \bar{C}_n$$

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

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

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

Utilisé lorsque l'on remplace $\boldsymbol{\Sigma}_C$ (connu) par $\hat{\boldsymbol{\Sigma}}_C$ 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 connaît $\mu_z = \mathbb{E}[Z_i]$, $\sigma_z^2 = \text{Var}[Z_i]$, 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 connaît $\mu_z = \mathbb{E}[Z_i]$, $\sigma_z^2 = \text{Var}[Z_i]$, etc.

Pour n répétitions, on a Z_1, \dots, 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 connaît $\mu_z = \mathbb{E}[Z_i]$, $\sigma_z^2 = \text{Var}[Z_i]$, etc.

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

Pour ajuster seulement **la moyenne**, remplacer Z_i par

$$\tilde{Z}_i = Z_i + \mu_z - \bar{Z}_n.$$

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 connaît $\mu_z = \mathbb{E}[Z_i]$, $\sigma_z^2 = \text{Var}[Z_i]$, etc.

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

Pour ajuster seulement **la moyenne**, remplacer Z_i par

$$\tilde{Z}_i = Z_i + \mu_z - \bar{Z}_n.$$

Pour ajuster **la moyenne et la variance**, remplacer Z_i par

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

où $S_{z,n}^2$ 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 $X_i = h(Z_i)$, $X_{\text{mm},i} = h(\tilde{Z}_i)$,

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

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

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

On peut montrer (exercice) que

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

quand $n \rightarrow \infty$, où $C^{(1)} = 1 - \sigma_z/S_{z,n}$, $C^{(2)} = \bar{Z}_n \sigma_z/S_{z,n} - \mu_z$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
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 $X_i = h(Z_i)$, $X_{\text{mm},i} = h(\tilde{Z}_i)$,

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

On peut montrer (exercice) que

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

quand $n \rightarrow \infty$, où $C^{(1)} = 1 - \sigma_z/S_{z,n}$, $C^{(2)} = \bar{Z}_n \sigma_z/S_{z,n} - \mu_z$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
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 $X_i = h(Z_i)$, $X_{\text{mm},i} = h(\tilde{Z}_i)$,

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

On peut montrer (exercice) que

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

quand $n \rightarrow \infty$, où $C^{(1)} = 1 - \sigma_z/S_{z,n}$, $C^{(2)} = \bar{Z}_n \sigma_z/S_{z,n} - \mu_z$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
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 $X_i = h(Z_i)$, $X_{\text{mm},i} = h(\tilde{Z}_i)$,

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

On peut montrer (exercice) que

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

quand $n \rightarrow \infty$, où $C^{(1)} = 1 - \sigma_z/S_{z,n}$, $C^{(2)} = \bar{Z}_n \sigma_z/S_{z,n} - \mu_z$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
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!

De plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(2)}] \neq 0$ en général.

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

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

On peut montrer (exercice) que

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

quand $n \rightarrow \infty$, où $C^{(1)} = 1 - \sigma_z/S_{z,n}$, $C^{(2)} = \bar{Z}_n \sigma_z/S_{z,n} - \mu_z$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
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!

De plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(2)}] \neq 0$ en général.

On pourrait utiliser les VCs $\bar{Z}_n - \mu_z$ et $S_{z,n}^2 - \sigma_z^2$ à la place.

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

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

On peut montrer (exercice) que

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

quand $n \rightarrow \infty$, où $C^{(1)} = 1 - \sigma_z/S_{z,n}$, $C^{(2)} = \bar{Z}_n \sigma_z/S_{z,n} - \mu_z$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
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!

De plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(2)}] \neq 0$ en général.

On pourrait utiliser les VCs $\bar{Z}_n - \mu_z$ et $S_{z,n}^2 - \sigma_z^2$ à la place.

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

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

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

On peut montrer (exercice) que

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

quand $n \rightarrow \infty$, où $C^{(1)} = 1 - \sigma_z/S_{z,n}$, $C^{(2)} = \bar{Z}_n \sigma_z/S_{z,n} - \mu_z$, $\beta_1 = \mathbb{E}[Z_i h'(Z_i)]$,
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!

De plus $\mathbb{E}[C^{(1)}] \neq 0$, $\mathbb{E}[C^{(2)}] \neq 0$ en général.

On pourrait utiliser les VCs $\bar{Z}_n - \mu_z$ et $S_{z,n}^2 - \sigma_z^2$ à la place.

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

On peut générer m copies indépendantes de cet estimateur.