

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.

On se restreint ici aux VCs linéaires.

Variables de contrôle (VC)

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

On se restreint ici aux VCs 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 $E[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

Variables de contrôle (VC)

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

On se restreint ici aux VCs 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 $E[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

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 $E[X_c] = E[X] = \mu$.

Variables de contrôle (VC)

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

On se restreint ici aux VCs 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 $E[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

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 $E[X_c] = E[X] = \mu$.

Comment choisir $\boldsymbol{\beta}$?

Variables de contrôle (VC)

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

On se restreint ici aux VCs 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 $E[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

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 $E[X_c] = E[X] = \mu$.

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.

On se restreint ici aux VCs 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 $E[\mathbf{C}] = \boldsymbol{\nu} = (\nu^{(1)}, \dots, \nu^{(q)})^t$.

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 $E[X_c] = E[X] = \mu$.

Comment choisir $\boldsymbol{\beta}$?

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

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

On a alors

$$\text{Var}[X_c] = \text{Var}[X] + \beta^t \Sigma_C \beta - 2\beta^t \Sigma_{CX}.$$

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

$$\text{Var}[X_c] = \text{Var}[X] + \beta^t \Sigma_C \beta - 2\beta^t \Sigma_{CX}.$$

Pour minimiser par rapport à β , on met le gradient par rapport à β égal à zéro:

$$0 = \nabla_{\beta} \text{Var}[X_c] = 2\Sigma_C \beta - 2\Sigma_{CX}.$$

Le minimum est donc atteint pour

$$\beta = \beta^* = \Sigma_C^{-1} \Sigma_{CX},$$

qui donne la variance minimale

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

où

$$R_{CX}^2 = \frac{\Sigma_{CX}^t \Sigma_C^{-1} \Sigma_{CX}}{\text{Var}[X]}$$

(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

$$\text{Var}[X_c] = \text{Var}[X] + \beta^t \Sigma_C \beta - 2\beta^t \Sigma_{CX}.$$

Pour minimiser par rapport à β , on met le gradient par rapport à β égal à zéro:

$$0 = \nabla_{\beta} \text{Var}[X_c] = 2\Sigma_C \beta - 2\Sigma_{CX}.$$

Le minimum est donc atteint pour

$$\beta = \beta^* = \Sigma_C^{-1} \Sigma_{CX},$$

qui donne la variance minimale

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

où

$$R_{CX}^2 = \frac{\Sigma_{CX}^t \Sigma_C^{-1} \Sigma_{CX}}{\text{Var}[X]}$$

(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 $\beta \neq \beta^*$, 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

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

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 $X^{(0)}, \dots, X^{(q)}$ 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

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

et

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

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

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

et

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $(1-R_{CX}^2)$ est réduite d'une fraction $\geq 1/(n-q-2)$.

Expériences pilotes pour estimer β^* ?

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:

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

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

Expériences pilotes pour estimer β^* ?

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:

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

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

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

Expériences pilotes pour estimer β^* ?

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:

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

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

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

Mais sous l'hypothèse de normalité,

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

C'est donc inefficace.