Composante du produit nº 12-001-X
Division des méthodes d'enquêtes auprès des entreprises

Article

Estimation par calage en utilisant l'inclinaison exponentielle dans les enquêtes par sondage

par Jae Kwang Kim

Décembre 2010

Statistique Canada Statistics Canada

Canada

Techniques d'enquête, décembre 2010
Vol. 36, N° 2, pp. 157-168
Statistique Canada, N° 12-001-X au catalogue

Estimation par calage en utilisant l'inclinaison exponentielle dans les enquêtes par sondage

Jae Kwang Kim [1]

Résumé

Nous considérons le problème de l'estimation des paramètres au moyen d'information auxiliaire, quand celle-ci prend la forme de moments connus. L'estimation par calage est un exemple type de l'utilisation des conditions des moments dans les enquêtes par sondage. Étant donné la forme paramétrique de la distribution originale des observations de l'échantillon, nous utilisons l'échantillonnage préférentiel avec distribution d'échantillonnage estimée de Henmi, Yoshida et Eguchi (2007) pour obtenir un estimateur amélioré. Si nous nous servons de la densité normale pour calculer les poids d'échantillonnage préférentiel, l'estimateur résultant prend la forme d'un estimateur par inclinaison exponentielle en une étape. Nous montrons que l'estimateur par inclinaison exponentielle proposé est asymptotiquement équivalent à l'estimateur par la régression, mais qu'il permet d'éviter les poids extrêmes et offre des avantages du point de vue des calculs par rapport à l'estimateur de la vraisemblance empirique. Nous discutons également de l'estimation de la variance et présentons les résultats d'une étude par simulation limitée.

Mots clés : Estimateur par étalonnage ; vraisemblance empirique ; calage au moyen de variables instrumentales ; échantillonnage préférentiel ; estimateur par la régression.

1. Introduction

Considérons le problème de l'estimation de $Y = \sum_{i=1}^{N} y_i$ pour une population finie de taille N. Soit A l'ensemble d'indices de l'échantillon obtenu selon un plan d'échantillonnage probabiliste. En plus de y_i, supposons que nous observons aussi un vecteur auxiliaire $\mathbf{x}_i$ de dimension p dans l'échantillon, tel que $\mathbf{X} = \sum_{i=1}^{N} \mathbf{x}_i$ est connu d'après une source externe. Nous voulons estimer Y en utilisant l'information auxiliaire $\mathbf{X}$.

L'estimateur de Horvitz-Thompson (HT) de la forme

$$\hat{Y}_d = \sum_{i \in A} d_i \, y_i, \tag{1}$$

où $d_i = 1/\pi_i$ est le poids d'échantillonnage et π_i est la probabilité d'inclusion de premier ordre, est sans biais pour Y. Toutefois, il n'utilise pas l'information fournie par $\mathbf{X}$. Selon Kott (2006), un estimateur par calage peut être défini comme l'estimateur de la forme

$$\hat{Y}_w = \sum_{i \in A} w_i \, y_i$$

où les poids w_i satisfont

$$\sum_{i \in A} w_i \, \mathbf{x}_i = \mathbf{X} \tag{2}$$

et $\hat{Y}_w$ est asymptotiquement sans biais par rapport au plan. L'estimation par calage est aujourd'hui très répandue dans les enquêtes par sondage, parce qu'elle assure la cohérence des résultats entre diverses enquêtes et améliore souvent l'efficacité (Särndal 2007).

L'estimateur par la régression, en utilisant les poids

$$w_i = d_i + (\mathbf{X} - \hat{\mathbf{X}}_d)' \left(\sum_{j \in A} d_j \, \mathbf{x}_j \, \mathbf{x}_j' \right)^{-1} d_i \, \mathbf{x}_i, \tag{3}$$

obtenus en minimisant

$$\sum_{i \in A} (w_i - d_i)^2 / d_i$$

sous la contrainte (2), est asymptotiquement sans biais par rapport au plan. Notons que, si un terme constant est inclus dans l'espace colonne de la matrice X, alors (2) implique que la taille de population N est connue. Si N est inconnu, on peut exiger que la somme des poids finaux soit égale à la somme des poids d'échantillonnage. Donc,

$$\sum_{i \in A} w_i = \hat{N}, \tag{4}$$

où

$$\hat{N} = \begin{cases} N & \text{si } N \text{ est connu} \\ \sum_{i \in A} d_i & \text{autrement,} \end{cases}$$

peut être imposé comme contrainte en plus de (2), ce qui donne les poids

$$w_i = \frac{\hat{N}}{\hat{N}_d} d_i + \left(\mathbf{X} - \frac{\hat{N}}{\hat{N}_d} \hat{\mathbf{X}}_d \right)' \left\{ \sum_{j \in A} d_j (\mathbf{x}_j - \bar{\mathbf{X}}_d) (\mathbf{x}_j - \bar{\mathbf{X}}_d)' \right\}^{-1} d_i (\mathbf{x}_i - \bar{\mathbf{X}}_d), \tag{5}$$

1. Jae Kwang Kim, Department of Statistics, Iowa State University, Ames, Iowa, 50011, États-Unis. Courriel : jkim@iastate.edu.

où $\hat{\mathbf{X}}_d = \sum_{i\in A} d_i \mathbf{x}_i$, $\hat{N}_d = \sum_{i\in A} d_i$ et $\bar{\mathbf{X}}_d = \hat{\mathbf{X}}_d / \hat{N}_d$. Nous définissons l'estimateur par la régression comme étant $\hat{Y}_{\text{reg}} = \sum_{i\in A} w_i y_i$ en utilisant les poids (5). L'estimateur par la régression peut être efficace si y_i est relié linéairement à $\mathbf{x}_i$ (Isaki et Fuller 1982 ; Fuller 2002), mais les poids dans cet estimateur peuvent prendre des valeurs négatives ou extrêmement grandes.

L'estimateur par calage par la vraisemblance empirique (EL pour *empirical likelihood*) dont discutent Chen et Qin (1993), Chen et Sitter (1999), Wu et Rao (2006), et Kim (2009) s'obtient en maximisant la pseudo-vraisemblance empirique

$$\sum_{i\in A} d_i \ln(w_i)$$

sous les contraintes (2) et (4). La solution du problème d'optimisation peut s'écrire

$$w_i = d_i \frac{1}{\lambda_0 + \boldsymbol{\lambda}_1' (\mathbf{x}_i - \mathbf{X}/\hat{N})}, \tag{6}$$

où λ_0 et $\boldsymbol{\lambda}_1$ satisfont les contraintes (2), (4) et $w_i > 0$ pour tout i. L'estimateur par calage EL est asymptotiquement équivalent à l'estimateur par la régression avec utilisation des poids (5) et évite l'obtention de poids négatifs si une solution existe, mais peut produire des poids extrêmement grands.

Comme la méthode de la vraisemblance empirique requiert la résolution d'équations non linéaires, les calculs peuvent être fastidieux. En outre, dans certains cas extrêmes, $\bar{\mathbf{X}} = N^{-1}\sum_{i=1}^N \mathbf{x}_i$ n'appartient pas à l'enveloppe convexe des $\mathbf{x}_i$ d'échantillon et la solution n'existe pas. Le cas échéant, la contrainte (2) peut être relâchée.

Rao et Singh (1997) ont résolu un problème similaire en permettant que

$$\left| \sum_{i\in A} w_i x_{ij} - X_j \right| \le \delta_j X_j, \quad j = 1, 2, \ldots, p,$$

pour un seuil de tolérance donné faible $\delta_j > 0$, où $X_j = \sum_{i=1}^N x_{ij}$. Notons que le choix $\delta_j = 0$ aboutit à la condition de calage exact (2). Rao et Singh (1997) ont choisi le seuil de tolérance δ_j en utilisant un facteur de rétrécissement dans la régression ridge, mais leur approche ne s'applique pas directement à la méthode de la vraisemblance empirique et le choix de δ_j n'est pas tout à fait clair. Chambers (1996), et Beaumont et Bocci (2008) ont également discuté de l'estimation par la régression ridge comme moyen d'éviter les poids extrêmes. Breidt, Claeskens et Opsomer (2005) ont suivi l'approche des splines pénalisées pour obtenir le calage ridge. Récemment, Park et Fuller (2009) ont élaboré une méthode d'obtention du facteur de rétrécissement δ_j en utilisant un modèle de régression en superpopulation avec composantes aléatoires.

Chen, Variyath et Abraham (2008) ont essayé de résoudre un problème semblable dans le contexte de la méthode du maximum de vraisemblance empirique et ont proposé une solution en ajoutant un point artificiel, tel que $\bar{\mathbf{X}} = N^{-1}\sum_{i=1}^N \mathbf{x}_i$ appartiendrait à l'enveloppe convexe des indices $\mathbf{x}_i$ augmentés. L'estimateur proposé dans Chen et coll. (2008) ne satisfait la propriété de calage qu'approximativement en ce sens que

$$\sum_{i\in A} w_i \mathbf{x}_i - \mathbf{X} = o_p(n^{-1/2}N). \tag{7}$$

Cette propriété de calage approximatif est intéressante, parce qu'elle permet une plus grande généralité dans le choix des poids. En particulier, quand la dimension de la variable auxiliaire $\mathbf{x}$ est grande, la contrainte de calage (2) peut être assez restreignante. Comme nous le montrons à la section 2, un estimateur satisfaisant la propriété de calage asymptotique (7) possède la plupart des propriétés désirables de l'estimateur par calage par la vraisemblance empirique et est efficace sur le plan des calculs.

Par le présent article, nous considérons une classe d'estimateurs de type vraisemblance empirique qui satisfont la propriété de calage approximatif (7). À la section 2, nous discutons de l'idée de l'échantillonnage préférentiel avec distribution d'échantillonnage estimée de Henmi et coll. (2007), et proposons un nouvel estimateur s'appuyant sur cette méthode. À la section 3, nous proposons une technique de troncature des poids pour éviter les poids de calage extrêmes. À la section 4, nous discutons de l'estimation de la variance de l'estimateur proposé. À la section 5, nous exposons les résultats d'une étude par simulation. Enfin, à la section 6, nous présentons nos conclusions.

2. Méthode proposée

Avant de présenter la méthode que nous proposons, nous discutons de l'échantillonnage préférentiel introduit par Henmi et coll. (2007). Supposons que $\mathbf{x}_i$ est observé dans toute la population, mais que y_i est observé uniquement dans l'échantillon. Nous émettons l'hypothèse d'un modèle de superpopulation pour $\mathbf{x}_i$ dont la densité $f(\mathbf{x}; \boldsymbol{\eta})$ est connue jusqu'à un paramètre $\boldsymbol{\eta} \in \Omega$. Le modèle de superpopulation caractérisé par la densité $f(\mathbf{x}; \boldsymbol{\eta})$ est un modèle de travail en ce sens qu'il est utilisé pour calculer un estimateur assisté par modèle (Särndal, Swenson et Wretman 1992).

Soit $\hat{\boldsymbol{\eta}}$ l'estimateur du pseudo-maximum de vraisemblance de $\boldsymbol{\eta}$ calculé d'après l'échantillon

$$\hat{\boldsymbol{\eta}} = \arg\max_{\Omega} \sum_{i\in A} d_i \ln\{f(\mathbf{x}_i; \boldsymbol{\eta})\}$$

et soit $\boldsymbol{\eta}_{0,N}$ l'estimateur du maximum de vraisemblance de $\boldsymbol{\eta}$ calculé d'après la population

$$\boldsymbol{\eta}_{0,N} = \arg\max_{\Omega} \sum_{i=1}^{N} \ln\{f(\mathbf{x}_i\,;\boldsymbol{\eta})\}.$$

En nous inspirant de Henmi et coll. (2007), nous pouvons construire le poids d'échantillonnage préférentiel estimé suivant

$$w_i = d_i \frac{f(\mathbf{x}_i\,;\boldsymbol{\eta}_{0N})}{f(\mathbf{x}_i\,;\hat{\boldsymbol{\eta}})}. \tag{8}$$

Afin de discuter des propriété asymptotiques de l'estimateur qui utilise les poids donnés par (8), supposons que nous avons une série de populations finies et d'échantillons, comme dans Isaki et Fuller (1982), tels que

$$\sum_{i\in A} d_i(\mathbf{x}_i', y_i)'\,(\mathbf{x}_i', y_i) - \sum_{i=1}^{N}(\mathbf{x}_i', y_i)'\,(\mathbf{x}_i', y_i) = O_p(n^{-1/2}N)$$

pour tout A possible et pour chaque N. Le théorème qui suit donne certaines propriétés asymptotiques de l'estimateur avec les poids d'échantillonnage préférentiel estimés donnés par (8).

Théorème 1. *Sous les conditions de régularité données à l'annexe A, l'estimateur* $\hat{Y}_w = \sum_{i\in A} w_i y_i$, *avec les* w_i *définis par* (8), *satisfait*

$$\sqrt{n}N^{-1}(\hat{Y}_w - \hat{Y}_l) = o_p(1), \tag{9}$$

où

$$\hat{Y}_l = \hat{Y}_d - \hat{\boldsymbol{\Sigma}}_{sy}'\,\hat{\boldsymbol{\Sigma}}_{ss}^{-1}\,\hat{\mathbf{S}}_{0d}, \tag{10}$$

$\hat{Y}_d$ *est défini en* (1), $\hat{\mathbf{S}}_{0d} = \sum_{i\in A} d_i \mathbf{s}_{i0}, \hat{\boldsymbol{\Sigma}}_{sy} = N^{-1}\sum_{i\in A} d_i\,\mathbf{s}_{i0}\,y_i$, *et* $\hat{\boldsymbol{\Sigma}}_{ss} = N^{-1}\sum_{i\in A} d_i\,\mathbf{s}_{i0}^{\otimes 2}$. *Ici,* $\mathbf{s}_{i0} = \partial \ln f(\mathbf{x}_i\,;\boldsymbol{\eta})/\partial\boldsymbol{\eta}\,|_{\boldsymbol{\eta}=\boldsymbol{\eta}_{0,N}}$ *et la notation* $B^{\otimes 2}$ *désigne* BB'.

La preuve du théorème 1 est présentée à l'annexe A. Comme $\mathbf{S}_{0N} \equiv \sum_{i=1}^{N}\mathbf{s}_{i0} = \mathbf{0}$, nous pouvons écrire (10) sous la forme

$$\hat{Y}_l = \hat{Y}_d + \hat{\boldsymbol{\Sigma}}_{sy}'\hat{\boldsymbol{\Sigma}}_{ss}^{-1}(\mathbf{S}_{0N} - \hat{\mathbf{S}}_{0d}),$$

qui est un estimateur par la régression de Y utilisant $\mathbf{s}_i(\boldsymbol{\eta}_{0N})$ comme variable auxiliaire. Par conséquent, sous des conditions de régularité, l'estimateur proposé utilisant l'échantillonnage préférentiel avec distribution estimée est asymptotiquement sans bais et possède une variance asymptotique qui n'est pas supérieure à celle de l'estimateur direct $\hat{Y}_d$. Notons que la validité du théorème 1 ne requiert pas que le modèle de travail $f(\mathbf{x}\,;\boldsymbol{\eta})$ soit vrai.

Si la densité de $\mathbf{x}_i$ est une densité normale multivariée, alors les poids donnés par (8) deviennent

$$w_i = d_i \frac{\phi(\mathbf{x}_i\,;\bar{\mathbf{X}}_N,\,\boldsymbol{\Sigma}_{xx,N})}{\phi(\mathbf{x}_i\,;\bar{\mathbf{X}}_d,\,\hat{\boldsymbol{\Sigma}}_{xx,d})}, \tag{11}$$

où $\bar{\mathbf{X}}_d$ est défini comme dans (5), $\hat{\boldsymbol{\Sigma}}_{xx,d} = \sum_{i\in A} d_i(\mathbf{x}_i - \bar{\mathbf{X}}_d)^{\otimes 2}/\hat{N}_d$, $\boldsymbol{\Sigma}_{xx,N} = \sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{X}}_N)^{\otimes 2}/N$ et $\phi(\mathbf{x}\,;\boldsymbol{\mu},\boldsymbol{\Sigma})$ est la densité de la loi normale multivariée de moyenne $\boldsymbol{\mu}$ et de matrice de variance-covariance $\boldsymbol{\Sigma}$. Si $\boldsymbol{\Sigma}_{xx,N}$ est inconnue et que seul $\bar{\mathbf{X}}_N$ est disponible, nous pouvons utiliser

$$w_i = d_i \frac{\phi(\mathbf{x}_i\,;\bar{\mathbf{X}}_N,\,\hat{\boldsymbol{\Sigma}}_{xx,d})}{\phi(\mathbf{x}_i\,;\bar{\mathbf{X}}_d,\,\hat{\boldsymbol{\Sigma}}_{xx,d})}. \tag{12}$$

Tillé (1998) a dérivé des poids similaires à ceux donnés par (12) dans le contexte des probabilités d'inclusion conditionnelles.

En général, le modèle paramétrique pour $\mathbf{x}_i$ est inconnu. Donc, nous considérons une approximation des poids d'échantillonnage préférentiel donnés par (8) en utilisant le critère d'information de Kullback-Leibler comme distance. Soit $f(\mathbf{x})$ une densité donnée pour $\mathbf{x}$ et soit P_0 l'ensemble des densités qui satisfont la contrainte de calage. Autrement dit,

$$P_0 = \left\{ f_0(\mathbf{x});\ \int f_0(\mathbf{x})\,d\mathbf{x} = 1,\ \int \mathbf{x} f_0(\mathbf{x})\,d\mathbf{x} = \bar{\mathbf{X}}_N \right\}.$$

Le problème d'optimisation en utilisant la distance de Kullback-Leibler peut s'exprimer sous la forme

$$\min_{f_0\in P_0} \int f_0(\mathbf{x}) \ln\left\{\frac{f_0(\mathbf{x})}{f(\mathbf{x})}\right\} d\mathbf{x}. \tag{13}$$

La solution de (13) est

$$f_0(\mathbf{x}) = f(\mathbf{x})\frac{\exp(\hat{\boldsymbol{\lambda}}'\mathbf{x})}{E\{\exp(\hat{\boldsymbol{\lambda}}'\mathbf{x})\}} \tag{14}$$

où $\hat{\boldsymbol{\lambda}}$ satisfait $\int \mathbf{x} f_0(\mathbf{x})\,d\mathbf{x} = \bar{\mathbf{X}}_N$. Donc, les poids d'échantillonnage préférentiel estimés donnés par (8) peuvent s'écrire, en utilisant la densité optimale (14), sous la forme

$$w_i = d_i \frac{f_0(\mathbf{x}_i)}{f(\mathbf{x}_i)} = d_i \exp(\hat{\lambda}_0 + \hat{\boldsymbol{\lambda}}_1'\mathbf{x}_i) \tag{15}$$

où $\hat{\lambda}_0$ et $\hat{\boldsymbol{\lambda}}_1$ satisfont les contraintes (2) et (4). Le déplacement de $f(\mathbf{x})$ et de $f_0(\mathbf{x})$ dans (14) est appelé inclinaison exponentielle (*exponential tilting*). Donc, un estimateur utilisant le poids (15) satisfaisant les contraintes de calage (2) et (4) peut être appelé estimateur par calage par inclinaison exponentielle (ET pour *exponential tilting*). Autrement dit, nous définissons l'estimateur par calage ET comme il suit

$$\hat{Y}_{\mathrm{ET}} = \sum_{i\in A} d_i \exp(\hat{\lambda}_0 + \hat{\boldsymbol{\lambda}}_1'\mathbf{x}_i)\,y_i, \tag{16}$$

où $\hat{\lambda}_0$ et $\hat{\boldsymbol{\lambda}}_1$ satisfont les contraintes (2) et (4). Les estimateurs basés sur l'inclinaison exponentielle ont été utilisés dans divers contextes. Consulter, par exemple, Efron (1981), Kitamura et Stutzer (1997), et Imbens (2002). Pour N connu, Folsom (1991) ainsi que Deville, Särndal et Sautory (1993) ont élaboré l'estimateur (16) en utilisant une approche fort différente.

Afin de calculer λ_0 et $\boldsymbol{\lambda}_1$ dans (16), à cause des contraintes de calage (2) et (4), nous devons résoudre les estimations suivantes :

$$\hat{U}_0(\boldsymbol{\lambda}) \equiv \sum_{i\in A} d_i \exp(\lambda_0 + \boldsymbol{\lambda}_1' \mathbf{x}_i) - \hat{N} = 0 \qquad (17)$$

$$\hat{\mathbf{U}}_1(\boldsymbol{\lambda}) \equiv \sum_{i\in A} d_i \exp(\lambda_0 + \boldsymbol{\lambda}_1'\mathbf{x}_i)\mathbf{x}_i - \mathbf{X} = \mathbf{0}, \qquad (18)$$

où $\boldsymbol{\lambda}' = (\lambda_0, \boldsymbol{\lambda}_1')$. En écrivant $\hat{\mathbf{U}}' = (\hat{U}_0, \hat{\mathbf{U}}_1')$, nous pouvons utiliser l'algorithme de type Newton de la forme

$$\hat{\boldsymbol{\lambda}}_{(t+1)} = \hat{\boldsymbol{\lambda}}_{(t)} - \left\{\frac{\partial}{\partial \boldsymbol{\lambda}'}\hat{\mathbf{U}}(\hat{\boldsymbol{\lambda}}_{(t)})\right\}^{-1} \hat{\mathbf{U}}(\hat{\boldsymbol{\lambda}}_{(t)})$$

et la solution peut s'écrire

$$\hat{\boldsymbol{\lambda}}_{1(t+1)} = \hat{\boldsymbol{\lambda}}_{1(t)} + \left\{\sum_{i\in A} w_{i(t)}(\mathbf{x}_i - \bar{\mathbf{X}}_{w(t)})^{\otimes 2}\right\}^{-1}\left(\mathbf{X} - \sum_{i\in A} w_{i(t)}\mathbf{x}_i\right), \qquad (19)$$

où $w_{i(t)} = d_i \exp(\hat{\lambda}_{0(t)} + \hat{\boldsymbol{\lambda}}_{1(t)}'\mathbf{x}_i)$ et $\bar{\mathbf{X}}_{w(t)} = \sum_{i\in A} w_{i(t)}\mathbf{x}_i / \sum_{i\in A} w_{i(t)}$, avec les valeurs initiales $\hat{\boldsymbol{\lambda}}_{1(0)} = \mathbf{0}$. Après avoir calculé $\hat{\boldsymbol{\lambda}}_{1(t)}$ en nous servant de (19), nous calculons $\hat{\lambda}_{0(t)}$ comme il suit

$$\exp(\hat{\lambda}_{0(t)}) = \frac{\hat{N}}{\sum_{i\in A} d_i \exp(\hat{\boldsymbol{\lambda}}_{1(t)}'\mathbf{x}_i)}. \qquad (20)$$

Notons que $w_{i(0)} = d_i\hat{N}/\hat{N}_d$ car $\hat{\boldsymbol{\lambda}}_{1(0)} = \mathbf{0}$. $\hat{\mathbf{U}}(\boldsymbol{\lambda})$ étant deux fois continuellement dérivable et convexe en $\boldsymbol{\lambda}$, la série $\hat{\boldsymbol{\lambda}}_{(t)}$ converge toujours si la solution de $\hat{\mathbf{U}}(\boldsymbol{\lambda}) = \mathbf{0}$ existe (Givens et Hoeting 2005). Le taux de convergence est quadratique en ce sens que

$$|\hat{\boldsymbol{\lambda}}_{1(t+1)} - \hat{\boldsymbol{\lambda}}_1| \le C|\hat{\boldsymbol{\lambda}}_{1(t)} - \hat{\boldsymbol{\lambda}}_1|^2$$

pour une constante C, où $\hat{\boldsymbol{\lambda}}_1 = \lim_{t\to\infty}\hat{\boldsymbol{\lambda}}_{1(t)}$.

Par construction, l'estimateur par inclinaison exponentielle (ET) en t étapes, défini par

$$\hat{Y}_{\mathrm{ET}(t)} = \sum_{i\in A} d_i \exp(\hat{\lambda}_{0(t)} + \hat{\boldsymbol{\lambda}}_{1(t)}'\mathbf{x}_i)\, y_i \qquad (21)$$

où $\hat{\lambda}_{0(t)}$ et $\hat{\boldsymbol{\lambda}}_{1(t)}$ sont calculés au moyen de (19) et (20), satisfait la contrainte de calage (2) pour une valeur suffisamment grande de t. En vertu de la forme récursive dans (19) avec $\hat{\boldsymbol{\lambda}}_{1(0)} = \mathbf{0}$, nous pouvons écrire

$$\hat{\boldsymbol{\lambda}}_{1(t)} = \sum_{j=0}^{t-1} (\mathbf{S}_{xx,w(j)})^{-1}(\tilde{\mathbf{X}}_N - \bar{\mathbf{X}}_{w(j)}), \qquad (22)$$

où $\tilde{\mathbf{X}}_N = \mathbf{X}/\hat{N}$ et $\mathbf{S}_{xx,w(j)} = \sum_{i\in A} w_{i(t)}(\mathbf{x}_i - \bar{\mathbf{X}}_{w(t)})^{\otimes 2}/\hat{N}$. Donc, l'estimateur ET en t étapes (21) peut s'écrire

$$\hat{Y}_{\mathrm{ET}(t)} = \hat{N}\frac{\sum_{i\in A} d_i g_{i(t)} y_i}{\sum_{i\in A} d_i g_{i(t)}},$$

où

$$g_{i(t)} = \prod_{j=0}^{t-1} \frac{\phi\left(\mathbf{x}_i; \tilde{\mathbf{X}}_N, \mathbf{S}_{xx,w(j)}\right)}{\phi\left(\mathbf{x}_i; \bar{\mathbf{X}}_{w(j)}, \mathbf{S}_{xx,w(j)}\right)}.$$

Le théorème qui suit présente certaines propriétés asymptotiques de l'estimateur par inclinaison exponentielle.

Théorème 2. *L'estimateur ET en t étapes* (21) *basé sur les équations* (19) *et* (20) *satisfait*

$$\sqrt{n}N^{-1}(\hat{Y}_{\mathrm{ET}(t)} - \hat{Y}_{\mathrm{reg}}) = o_p(1), \qquad (23)$$

pour chaque $t = 1, 2, \ldots$, où $\hat{Y}_{\mathrm{reg}}$ est l'estimateur par la régression en utilisant les poids de régression (5).

La preuve du théorème 2 est présentée à l'annexe B. Le théorème 2 donne l'équivalence asymptotique entre l'estimateur ET en t étapes et l'estimateur par la régression. Contrairement à ce dernier, les poids de l'estimateur ET sont toujours positifs. Pour une valeur suffisamment grande de t, l'estimateur ET en t étapes satisfait la contrainte de calage (2). Deville et Särndal (1992) ont prouvé le résultat (23) pour le cas particulier où $t \to \infty$.

Remarque 1. *L'estimateur ET en une étape, défini par $\hat{Y}_{\mathrm{ET}(1)}$, possède un paramètre d'inclinaison ayant une expression analytique*

$$\hat{\boldsymbol{\lambda}}_{1(1)} = \left\{\sum_{i\in A} d_i(\mathbf{x}_i - \bar{\mathbf{X}}_d)^{\otimes 2} \Big/ \hat{N}_d\right\}^{-1}(\tilde{\mathbf{X}}_N - \bar{\mathbf{X}}_d), \qquad (24)$$

où $\tilde{\mathbf{X}}_N = \mathbf{X}/\hat{N}$ et $\bar{\mathbf{X}}_d = \sum_{i\in A} d_i\mathbf{x}_i / \sum_{i\in A} d_i$. En vertu du théorème 2, *l'estimateur ET en une étape est asymptotiquement équivalent à l'estimateur par la régression, mais la contrainte de calage* (2) *n'est pas nécessairement satisfaite. En utilisant le théorème* 2 *appliqué à $\mathbf{x}_i$ au lieu de y_i, on peut montrer que l'estimateur ET en une étape satisfait la contrainte de calage approximatif décrite en* (7).

Remarque 2. *L'estimateur ET peut également être dérivé en trouvant les poids qui minimisent*

$$Q(w) = \sum_{i \in A} w_i \ln\left(\frac{w_i}{d_i}\right) \qquad (25)$$

sous les contraintes (2) *et* (4). *La fonction objectif* (25) *est souvent appelée fonction de discrimination minimale. La valeur minimale de* $Q(w)$ *est zéro si* (4) *est la seule contrainte de calage et elle croît de manière monotone si des contraintes de calage additionnelles sont imposées.*

3. Calage au moyen de variables instrumentales

Nous considérons une extension de la méthode proposée à la section 2 à une classe plus générale d'estimateurs par calage ET utilisant des variables instrumentales. Estevao et Särndal (2000) ainsi que Kott (2003) ont discuté de l'utilisation de variables instrumentales pour l'estimation par calage dans le contexte de simulations limitées. Soit $\mathbf{z}_i = \mathbf{z}(\mathbf{x}_i)$ une variable instrumentale dérivée de $\mathbf{x}_i$, où la fonction $\mathbf{z}(\cdot)$ doit être déterminée. L'estimateur par inclinaison exponentielle avec variable instrumentale (IVET) (pour *instrumental-variable exponential tilting*) en se servant de la variable instrumentale $\mathbf{z}_i$ peut être défini comme

$$\hat{Y}_{\text{IVET}} = \sum_{i \in A} w_i y_i = \sum_{i \in A} d_i \exp(\hat{\lambda}_0 + \hat{\boldsymbol{\lambda}}_1' \mathbf{z}_i)\, y_i, \qquad (26)$$

où $\hat{\lambda}_0$ et $\hat{\boldsymbol{\lambda}}_1$ sont calculé d'après (2) et (4). Notons que l'estimateur IVET donné par (26) appartient à une classe d'estimateurs indicés par $\mathbf{z}_i$. L'approche de la variable instrumentale définie en (26) offre plus de souplesse pour créer l'estimateur ET. Le choix de $\mathbf{z}_i = \mathbf{x}_i$ aboutit à l'estimateur ET standard donné par (16), mais une certaine transformation $\mathbf{z}_i = \mathbf{z}(\mathbf{x}_i)$ peut rendre l'estimateur ET donné par (26) plus intéressant en pratique. La solution des équations de calage peut être obtenue itérativement comme il suit

$$\hat{\boldsymbol{\lambda}}_{1(t+1)} = \hat{\boldsymbol{\lambda}}_{1(t)} + \left\{\sum_{i \in A} w_{i(t)} (\mathbf{x}_i - \overline{\mathbf{X}}_{w(t)})(\mathbf{z}_i - \overline{\mathbf{Z}}_{w(t)})'\right\}^{-1} \left(\mathbf{X} - \sum_{i \in A} w_{i(t)} \mathbf{x}_i\right), \qquad (27)$$

où $w_{i(t)} = d_i \exp(\hat{\lambda}_{0(t)} + \hat{\boldsymbol{\lambda}}_{1(t)}' \mathbf{z}_i)$ et $\overline{\mathbf{Z}}_{w(t)} = \sum_{i \in A} w_{i(t)} \mathbf{z}_i / \sum_{i \in A} w_{i(t)}$, avec l'équation (20) inchangée et $\hat{\boldsymbol{\lambda}}_{1(0)} = \mathbf{0}$.

L'estimateur IVET (26) est utilité pour créer des poids finaux ayant des valeurs moins extrêmes. Puisque dans (26), le poids final est une fonction de $\mathbf{z}_i$, nous pouvons borner $g_i = w_i / d_i$ en donnant des bornes à $\mathbf{z}_i$. Pour créer la variable $\mathbf{z}_i$ bornée, nous pouvons utiliser une version tronquée de $\mathbf{x}_i$, désignée par $\mathbf{z}_i = (z_{i1}, z_{i2}, \ldots, z_{ip})$, où

$$z_{ij} = \begin{cases} x_{ij} & \text{si } |x_{ij} - \overline{x}_j| \le C_j S_j \\ \overline{x}_j + C_j S_j & \text{si } x_{ij} > \overline{x}_j + C_j S_j \\ \overline{x}_j - C_j S_j & \text{si } x_{ij} < \overline{x}_j - C_j S_j, \end{cases} \qquad (28)$$

$\overline{x}_j = N^{-1} \sum_{i \in A} d_i x_{ij}$, $S_j^2 = N^{-1} \sum_{i \in A} d_i (x_{ij} - \overline{x}_j)^2$, et C_j est un seuil pour la détection des valeurs aberrantes, par exemple, $C_j = 3$. Donc, l'estimateur IVET utilisant la variable instrumentale obtenue en tronquant $\mathbf{x}_i$ peut être utilisé comme alternative à la troncature des poids.

Au lieu d'employer la variable instrumentale tronquée $\mathbf{z}_i$ dans (28), nous pouvons considérer la variable instrumentale suivante

$$\mathbf{z}_i = \mathbf{x}_i \Phi_i$$

pour une matrice symétrique Φ_i telle que $\mathbf{z}_i$ est bornée. Un choix appropriée de Φ_i peut aussi améliorer l'efficacité de l'estimateur IVET résultant. En effet, en nous appuyant sur le même argument découlant du théorème 2, nous voyons que l'estimateur ET avec variable instrumentale (26) utilisé avec les équations (20) et (27) est asymptotiquement équivalent à

$$\hat{Y}_{\text{IV, reg}} = \tilde{Y}_d + (\mathbf{X} - \tilde{\mathbf{X}}_d)' \hat{\mathbf{B}}_z \qquad (29)$$

où

$$(\tilde{\mathbf{X}}_d', \tilde{Y}_d) = \left(\frac{\hat{N}}{\hat{N}_d}\right)(\hat{\mathbf{X}}_d', \hat{Y}_d)$$

et

$$\hat{\mathbf{B}}_z = \left\{\sum_{i \in A} d_i (\mathbf{z}_i - \overline{\mathbf{Z}}_d)(\mathbf{x}_i - \overline{\mathbf{X}}_d)'\right\}^{-1} \sum_{i \in A} d_i (\mathbf{z}_i - \overline{\mathbf{Z}}_d) y_i. \qquad (30)$$

L'estimateur (29) prend la forme d'un estimateur par la régression que nous dénommons estimateur par la régression à variable instrumentale. Donc, sous le choix $\mathbf{z}_i = \Phi_i \mathbf{x}_i$, l'estimateur par la régression à variable instrumentale peut s'écrire comme (29) avec

$$\hat{\mathbf{B}}_z = \left\{\sum_{i \in A} d_i (\mathbf{x}_i - \overline{\mathbf{X}}_d) \Phi_i (\mathbf{x}_i - \overline{\mathbf{X}}_d)'\right\}^{-1} \sum_{i \in A} d_i (\mathbf{x}_i - \overline{\mathbf{X}}_d) \Phi_i\, y_i$$

et sa variance est minimisée pour $\Phi_i = V_i^{-1}$ où V_i est la variance de y_i sous le modèle sachant $\mathbf{x}_i$ (Fuller 2009). Ici, la variance sous le modèle est la variance sous le modèle de superpopulation de travail pour la régression de y_i sur $\mathbf{x}_i$. Donc, nous pouvons utiliser la variable instrumentale pour accroître l'efficacité de l'estimateur par calage résultant, en plus d'éviter des poids finaux extrêmes. De surcroît, nous pouvons tronquer la variable instrumentale optimale comme dans (28) afin de borner les poids finaux. Une étude plus approfondie du choix optimal de Φ dépasse le cadre du présent article et sera le sujet de travaux de recherche à venir.

Remarque 3. *Deville et Särndal* (1992) *ont également considéré des poids de calage à étendue restreinte de la forme*

$$w_i = d_i g_i(\hat{\boldsymbol{\lambda}}) = d_i \frac{L(U-1)+U(1-L)\exp(K\hat{\boldsymbol{\lambda}}'\mathbf{x}_i)}{(U-1)+(1-L)\exp(K\hat{\boldsymbol{\lambda}}'\mathbf{x}_i)}, \quad (31)$$

où $K=(U-L)/\{(1-L)(U-1)\}$, *pour des valeurs de* L *et* U *telles que* $0<L<1<U$. *Si les contraintes de calage* (2) *et* (4) *doivent être satisfaites, nous pouvons utiliser* $\hat{\lambda}_0 + \hat{\boldsymbol{\lambda}}_1'\mathbf{x}_i$ *au lieu de* $\hat{\boldsymbol{\lambda}}'\mathbf{x}_i$ *dans* (31). *L'estimateur par calage résultant est asymptotiquement équivalent à l'estimateur par la régression avec utilisation des poids donnés par* (5), *tandis que l'estimateur IVET est asymptotiquement équivalent à l'estimateur par la régression à variable instrumentale* (29). *Les calculs en vue d'obtenir* $\hat{\boldsymbol{\lambda}}$ *sont assez compliqués, parce que* $\partial g_i(\boldsymbol{\lambda})/\partial\boldsymbol{\lambda}$ *n'est pas facile à évaluer dans* (31). *Dans l'estimateur IVET, le calcul, donné par* (27), *est simple.*

Afin de comparer la pondération proposée aux méthodes existantes, considérons l'exemple artificiel d'un échantillon aléatoire simple de taille $n=5$ où $x_k=k$, $k=1, 2, ..., 5$. Nous exécutons les calculs pour trois moyennes de population de x ; $\bar{X}_N=3$, $\bar{X}_N=4{,}5$, et $\bar{X}_N=6$. Le tableau 1 donne les poids résultants pour l'estimateur par la régression, l'estimateur par la vraisemblance empirique (EL), l'estimateur par inclinaison exponentielle (ET) en t étapes (16) avec $t=1$ et $t=10$, ainsi que l'estimateur par inclinaison exponentielle à variable instrumentale (IVET) en t étapes (26) avec $t=1$ et $t=10$. Pour l'estimateur IVET, nous créons la variable instrumentale z_i telle que

$$z_i=\begin{cases}1{,}5 & \text{si } x_i \le 1{,}5\\ x_i & \text{si } x_i \in (1{,}5\,;\ 4{,}5)\\ 4{,}5 & \text{si } x_i \ge 4{,}5.\end{cases}$$

La dernière colonne du tableau 1 donne la moyenne estimée de X en utilisant les poids de calage respectifs. Tous les poids sont égaux à $1/n=0{,}2$ pour $\bar{X}_N=3$. L'estimateur par la régression est linéairement croissant en x_i, mais possède des poids négatifs pour les populations de moyenne $\bar{X}_N=4{,}5$ et de moyenne $\bar{X}_N=6$. Pour la population de moyenne $\bar{X}_N=6$, les poids n'ont pas pu être calculés pour la méthode EL, parce que $\bar{X}_N$ se situe en dehors de l'intervalle des valeurs x_i d'échantillon. Dans ce cas extrême où $\bar{X}_N=6$, la méthode ET fournit des poids non négatifs en sacrifiant la contrainte de calage et l'estimateur EL possède des poids plus extrêmes que l'estimateur ET ou que l'estimateur IVET en ce sens que le poids pour $k=5$ est le plus grand observé parmi les estimateurs étudiés. Le poids pour l'estimateur ET en une étape est proche de celui de l'estimateur par la régression pour une grande valeur de x_i, mais il est proche de celui de l'estimateur EL pour une petite valeur de x_i. Les estimateurs ET en 10 étapes ont de meilleures propriétés de calage en ce sens que l'erreur quadratique, $(\sum_{k=1}^5 w_k x_k - \bar{X}_N)^2$, est plus faible que pour l'estimateur ET en une étape. L'estimateur ET et l'estimateur IVET donnent presque les mêmes estimations de $\bar{X}_N$ pour les deux valeurs de t, mais l'estimateur IVET produit des poids moins extrêmes que l'estimateur ET.

4. Estimation de la variance

Examinons maintenant l'estimation de la variance des estimateurs par calage ET des sections 2 et 3. Comme les paramètres estimés $(\hat{\lambda}_0, \hat{\boldsymbol{\lambda}}_1')$ qui figurent dans l'estimateur par calage ET (16) ont une certaine variabilité d'échantillonnage, la méthode d'estimation de la variance doit en tenir compte. Dans ce cas, l'estimation de la variance peut souvent être obtenue par une méthode de linéarisation ou par une méthode de rééchantillonnage, ou réplication (Wolter 2007). Pour la discussion de la méthode de linéarisation, posons que la variance de l'estimateur HT donné en (1) est estimée de manière convergente par

$$\hat{V}(\hat{Y}_d)=\sum_{i\in A}\sum_{j\in A}\Omega_{ij}y_i y_j. \quad (32)$$

Pour l'estimateur ET, l'estimateur de variance par linéarisation peut être obtenu au moyen de la formule d'estimation de la variance par linéarisation établie pour l'estimateur par la régression, comme dans Deville et Särndal (1992), en se servant de l'équivalence asymptotique entre l'estimateur par calage ET et l'estimateur par la régression montrée dans le théorème 2. En particulier, si l'on connait la taille de population N, un estimateur de variance par linéarisation pour l'estimateur IVET (26) peut s'écrire

$$\hat{V}(\hat{Y}_{\text{IVET}})=\sum_{i\in A}\sum_{j\in A}\Omega_{ij}\,g_i\,g_j\,\hat{e}_i\,\hat{e}_j \quad (33)$$

où Ω_{ij} correspond aux coefficients de l'estimateur de variance (32), $g_i=w_i/d_i$ est le facteur d'ajustement des poids et $\hat{e}_i=y_i-\bar{Y}_d-(\mathbf{x}_i-\bar{\mathbf{X}}_d)'\hat{\mathbf{B}}_z$, où $\hat{\mathbf{B}}_z$ est défini dans (30). Le choix $\mathbf{z}_i=\mathbf{x}_i$ dans (33) produit l'estimateur de variance linéarisé pour l'estimateur ET donné par (16). La démonstration de la convergence de l'estimateur de variance (33) peut être consultée dans Kim et Park (2010).

Pour l'estimateur ET en une étape, il est facile de mettre en œuvre une méthode de rééchantillonnage (réplication). Soit l'estimateur de variance par rééchantillonnage de la forme

$$\hat{V}_{\text{rep}}=\sum_{k=1}^{L}c_k(\hat{Y}_d^{(k)}-\hat{Y}_d)^2, \quad (34)$$

où L est le nombre de répliques, et c_k est le facteur de réplication associé à la réplique k, $\hat{Y}_d^{(k)}=\sum_{i\in A}d_i^{(k)}y_i$, et $d_i^{(k)}$ est la k^{e} réplique du poids d'échantillonnage d_i. Par exemple, l'estimateur de variance par rééchantillonnage (34) inclut le jackknife et le bootstrap (voir Rust et Rao 1996). Supposons que l'estimateur de variance par rééchantillonnage (34) est un estimateur convergent pour la variance de $\hat{Y}_d$. La k^{e} réplique de l'estimateur ET en une étape peut être calculée par

$$\hat{Y}_{\mathrm{ET}(1)}^{(k)} = \sum_{i\in A} d_i^{(k)} \exp(\hat{\lambda}_{0(1)}^{(k)} + \hat{\boldsymbol{\lambda}}_{1(1)}^{(k)'}\mathbf{z}_i)\, y_i \qquad (35)$$

où

$$\hat{\boldsymbol{\lambda}}_{1(1)}^{(k)} = \left\{\sum_{i\in A} d_i^{(k)}(\mathbf{x}_i - \bar{\mathbf{X}}_d^{(k)})\,(\mathbf{z}_i - \bar{\mathbf{Z}}_d^{(k)})' \Big/ \hat{N}_d^{(k)}\right\}^{-1} (\mathbf{X}/\hat{N}^{(k)} - \bar{\mathbf{X}}_d^{(k)}),$$

$$\hat{N}^{(k)} = \begin{cases} N & \text{si } \hat{N} = N \\ \hat{N}_d^{(k)} = \sum_{i\in A} d_i^{(k)} & \text{si } \hat{N} = \hat{N}_d, \end{cases}$$

$$(\bar{\mathbf{X}}_d^{(k)},\ \bar{\mathbf{Z}}_d^{(k)}) = \frac{\sum_{i\in A} d_i^{(k)}(\mathbf{x}_i,\ \mathbf{z}_i)}{\sum_{i\in A} d_i^{(k)}},$$

et

$$\exp(\hat{\lambda}_{0(1)}^{(k)}) = \frac{\hat{N}}{\sum_{i\in A} d_i^{(k)} \exp(\mathbf{z}'\hat{\boldsymbol{\lambda}}_{1(1)}^{(k)})}.$$

L'estimateur de variance par rééchantillonnage défini par

$$\hat{V}_{\mathrm{rep}} = \sum_{k=1}^{L} c_k (\hat{Y}_{\mathrm{ET}}^{(k)} - \hat{Y}_{\mathrm{ET}})^2, \qquad (36)$$

où $\hat{Y}_{\mathrm{ET}}^{(k)}$ est défini dans (35), peut être utilisé pour estimer la variance de l'estimateur par calage ET donné par (26).

5. Étude par simulation

Afin d'étudier les propriétés en échantillon fini des estimateurs proposés, nous avons effectué une étude par simulation limitée. Dans la simulation, nous avons généré indépendamment deux populations finies de taille $N = 10\,000$. Pour la population A, la population finie est générée en partant d'une population infinie spécifiée par $x_i \sim \exp(1) + 1$; $y_i = 3 + x_i + x_i e_i$, $e_i \,|\, x_i \sim N(0,1)$; $z_i \,|\, (x_i, y_i) \sim \chi^2(1) + |y_i|$. Dans le cas de la population B, les $(x_i,\ e_i,\ z_i)$ sont les mêmes que dans la population A, mais $y_i = (5 - 1/\sqrt{8}) + 1/\sqrt{8}(x_i - 2)^2 + e_i$. La variable auxiliaire x_i, est utilisée pour le calage et z_i est la mesure de taille utilisée pour l'échantillonnage avec probabilités inégales. À partir des deux populations finies ainsi obtenues, nous avons généré indépendamment $M = 10\,000$ échantillons Monte Carlo de taille n sous les deux plans d'échantillonnage décrits plus loin. Le paramètre d'intérêt est la moyenne de population de y et nous supposons que la taille de population N est connue.

Les conditions de simulation peuvent être décrites comme un plan factoriel $2 \times 2 \times 8 \times 2$ à quatre facteurs, soit a) deux types de population finie, b) le mécanisme d'échantillonnage : échantillonnage aléatoire simple et échantillonnage avec probabilité proportionnelle à la taille (z_i) avec remise, c) la méthode de calage : pas de calage, estimateur par la régression, méthode EL donnée par (6) avec $t = 1$ et $t = 10$, méthode ET en t étapes donnée par (21) avec $t = 1$ et $t = 10$, et méthode IVET donnée par (26) avec $t = 1$ et $t = 10$, et d) taille de l'échantillon : $n = 100$ et $n = 200$. Puisque l'on suppose que N est connu, les estimateurs par calage sont calculés de façon à satisfaire $\sum_{i=1}^{n} w_i(1,\ x_i) = (1,\ \bar{X}_N)$ dans les deux populations. Pour la méthode IVET (26), la variable instrumentale z_i est créée en utilisant les définitions données en (28) avec le seuil $C = 3$.

Tableau 1
Exemple de poids de calage avec un échantillon de taille $n = 5$

Méthode	$\bar{X}_N$	x_i 1	2	3	4	5	$\hat{X}_N$
Reg.	3,0	0,200	0,200	0,200	0,200	0,200	3,0
	4,5	-0,100	0,050	0,200	0,035	0,500	4,5
	6,0	-0,400	-0,100	0,200	0,500	0,800	6,0
EL	3,0	0,200	0,200	0,200	0,200	0,200	3,0
	4,5	0,033	0,043	0,063	0,115	0,746	4,5
	6,0	S.O.	S.O.	S.O.	S.O.	S.O.	S.O.
ET ($t = 1$)	3,0	0,200	0,200	0,200	0,200	0,200	3,0
	4,5	0,027	0,057	0,100	0,255	0,540	4,2
	6,0	0,002	0,009	0,039	0,173	0,777	4,7
ET ($t = 10$)	3,0	0,200	0,200	0,200	0,200	0,200	3,0
	4,5	0,009	0,027	0,078	0,227	0,659	4,5
	6,0	0,000	0,000	0,000	0,001	0,999	5,0
IVET ($t = 1$)	3,0	0,200	0,200	0,200	0,200	0,200	3,0
	4,5	0,030	0,047	0,121	0,309	0,493	4,2
	6,0	0,003	0,006	0,041	0,267	0,683	4,6
IVET ($t = 10$)	3,0	0,200	0,200	0,200	0,200	0,200	3,0
	4,5	0,007	0,015	0,066	0,294	0,618	4,5
	6,0	0,000	0,000	0,000	0,087	0,913	4,9

Rég., estimateur par la régression ; EL, vraisemblance empirique (*empirical likelihood*) ; ET, inclinaison exponentielle (*exponential tilting*) ; IVET, inclinaison exponentielle avec variable instrumentale (*instrumental variable exponential tilting*) ; S.O., sans objet.

En utilisant les échantillons Monte Carlo produits comme il est mentionné plus haut, nous avons calculé les biais et les erreurs quadratiques moyennes de huit estimateurs de la moyenne de population de y, la variable d'intérêt. Les résultats sont présentés au tableau 2. Les estimateurs par calage contiennent un biais, mais celui-ci est faible si le modèle de régression est vérifié ou si la taille d'échantillon est grande. Dans la population A, le modèle de régression linéaire est vérifié et l'estimateur par la régression est efficace selon les erreurs quadratiques moyennes. Par contre, il ne l'est pas dans la population B, parce que le modèle utilisé n'est pas bien ajusté. Les sept estimateurs par calage ont des propriétés comparables pour la taille d'échantillon la plus grande. L'estimateur IVET en dix étapes donne d'aussi bons résultats que l'estimateur par la régression dans la population A et possède des propriétés un peu meilleures que les six autres estimateurs par calage. Dans la population B, l'estimateur IVET en dix étapes est, de tous les estimateurs par calage étudiés, celui dont la performance est la meilleure.

En plus de l'estimation ponctuelle, nous avons examiné l'estimation de la variance. Nous n'avons considéré l'estimation de la variance que pour les estimateurs ET et IVET en t étapes. Nous avons calculé l'estimateur de variance par linéarisation (33) et l'estimateur de variance par rééchantillonnage (36) pour chaque estimateur dans chaque échantillon. Dans la méthode de rééchantillonnage, nous avons utilisé la méthode du jackknife avec suppression d'une unité pour chaque réplique. Nous avons calculé les biais relatifs des estimateurs de variance en divisant le biais Monte Carlo de l'estimateur de variance par la variance Monte Carlo. Les biais relatifs Monte Carlo des estimateurs de variance par linéarisation et des estimateurs de variance par rééchantillonnage sont présentés au tableau 3. Le biais relatif théorique des estimateurs de variance est d'ordre $o(1)$, ce qui concorde avec les résultats des simulations présentés au tableau 3. L'estimateur de variance par linéarisation sous-estime légèrement la variance réelle, parce qu'il omet le terme de deuxième ordre dans la linéarisation de Taylor. L'estimateur de variance par rééchantillonnage présente un léger biais positif dans la simulation. Les biais des estimateurs de variance sont généralement plus faibles en valeur absolue dans la population A, parce que le modèle linéaire est vérifié. Dans la population B, le biais des estimateurs de variance est moins important pour l'estimateur IVET que pour l'estimateur ET, car des poids moins extrême sont utilisés dans l'estimateur IVET.

6. Conclusion

Nous avons examiné le problème de l'estimation de Y en nous servant d'information auxiliaire de la forme $E\{U(\mathbf{X})\} = 0$ avec une fonction connue $U(\cdot)$. Nous avons considéré la classe des estimateurs linéaires de la forme $\hat{Y} = \sum_{i\in A} w_i y_i$ avec $\sum_{i\in A} w_i \{1,\, U(\mathbf{x}_i)\} = (\hat{N},\, 0)$ et $w_i > 0$ Si la densité $f(\mathbf{x}\,;\boldsymbol{\eta})$ de X est connue jusqu'à $\boldsymbol{\eta} \in \Omega$, nous pouvons mettre en oeuvre une estimation efficace en utilisant le poids d'échantillonnage préférentiel estimé

$$w_i \propto d_i \frac{f(x_i\,;\,\boldsymbol{\eta}_{0,N})}{f(x_i\,;\hat{\boldsymbol{\eta}})},$$

où d_i représente les poids initiaux, et où $\boldsymbol{\eta}_{0,N}$ et $\hat{\boldsymbol{\eta}}$ sont les estimateurs du maximum de vraisemblance de $\boldsymbol{\eta}$ basé sur la population et sur l'échantillon, respectivement. Si la forme paramétrique de $f(\mathbf{x}\,;\boldsymbol{\eta})$ est inconnue, nous pouvons utiliser les poids transformés par inclinaison exponentielle de la forme

$$w_{i(\lambda)} \propto \exp\{\boldsymbol{\lambda}' U(\mathbf{x}_i)\},$$

où $\boldsymbol{\lambda}$ est déterminé en vue de satisfaire

$$\sum_{i\in A} w_{i(\lambda)} U(\mathbf{x}_i) = 0. \tag{37}$$

Tableau 2
Biais Monte Carlo et erreurs quadratiques moyenne Monte Carlo des estimateurs ponctuels de la moyenne de y, basés sur 10 000 échantillons Monte Carlo

Population	Taille de l'échantillon	Estimateur	EAS		PPT	
			Biais	EQM	Biais	EQM
A	100	Pas de calage	0,00	0,02398	0,00	0,02023
		Estimateur par régression	0,00	0,01261	0,00	0,01289
		Estimateur EL (t = 1)	0,01	0,01369	0,01	0,01353
		Estimateur EL (t = 10)	0,00	0,01285	0,00	0,01289
		Estimateur ET (t = 1)	0,01	0,01334	0,01	0,01353
		Estimateur ET (t = 10)	0,00	0,01269	0,00	0,01289
		Estimateur IVET (t = 1)	0,01	0,01309	0,01	0,01330
		Estimateur IVET (t = 10)	0,00	0,01263	0,00	0,01289
	200	Pas de calage	0,00	0,01069	0,00	0,00925
		Estimateur par régression	0,00	0,00595	0,00	0,00568
		Estimateur EL (t = 1)	0,01	0,00632	0,01	0,00604
		Estimateur EL (t = 10)	0,00	0,00597	0,00	0,00568
		Estimateur ET (t = 1)	0,00	0,00616	0,01	0,00578
		Estimateur ET (t = 10)	0,00	0,00596	0,00	0,00568
		Estimateur IVET (t = 1)	0,00	0,00605	0,01	0,00574
		Estimateur IVET (t = 10)	0,00	0,00591	0,00	0,00567

Tableau 2 (suite)
Biais Monte Carlo et erreurs quadratiques moyenne Monte Carlo des estimateurs ponctuels de la moyenne de *y*, basés sur 10 000 échantillons Monte Carlo

Population	Taille de l'échantillon	Estimateur	EAS		PPT	
			Biais	EQM	Biais	EQM
B	100	Pas de calage	0,00	0,02044	0,00	0,01692
		Estimateur par régression	-0,01	0,01473	0,00	0,01461
		Estimateur EL ($t = 1$)	0,01	0,01652	0,01	0,01516
		Estimateur EL ($t = 10$)	0,00	0,01490	0,01	0,01472
		Estimateur ET ($t = 1$)	0,00	0,01516	0,01	0,01483
		Estimateur ET ($t = 10$)	0,00	0,01470	0,00	0,01459
		Estimateur IVET ($t = 1$)	0,00	0,01497	0,00	0,01458
		Estimateur IVET ($t = 10$)	0,00	0,01472	0,00	0,01453
	200	Pas de calage	0,00	0,00888	0,00	0,00823
		Estimateur par régression	-0,01	0,00705	0,00	0,00735
		Estimateur EL ($t = 1$)	0,01	0,00769	0,01	0,00764
		Estimateur EL ($t = 10$)	0,00	0,00715	0,01	0,00745
		Estimateur ET ($t = 1$)	0,00	0,00723	0,01	0,00749
		Estimateur ET ($t = 10$)	0,00	0,00706	0,01	0,00734
		Estimateur IVET ($t = 1$)	0,00	0,00704	0,00	0,00728
		Estimateur IVET ($t = 10$)	0,00	0,00699	0,00	0,00725

EAS, échantillonnage aléatoire simple ; PPT, échantillonnage avec probabilité proportionnelle à la taille ; EQM, erreur quadratique moyenne ; EL, vraisemblance empirique (*empirical likelihood*) ; ET, inclinaison exponentielle (*exponential tilting*) ; IVET, inclinaison exponentielle avec variable instrumentale (*instrumental-variable exponential tilting*).

Tableau 3
Biais relatifs Monte Carlo des estimateurs de variance, basés sur 10 000 échantillons Monte Carlo

Population	Taille de l'échantillon	Estimateur	Linéarisation		Rééchantillonnage	
			EAS	PPT	EAS	PPT
A	100	ET ($t = 1$)	-7,02	-2,66	10,65	4,11
		ET ($t = 10$)	-4,91	-0,80	5,60	0,67
		IVET ($t = 1$)	-5,28	-3,63	7,67	2,25
		IVET ($t = 10$)	-4,11	-0,87	4,96	0,41
	200	ET ($t = 1$)	-3,97	-0,19	3,65	0,57
		ET ($t = 10$)	-2,93	0,87	2,23	-0,35
		IVET ($t = 1$)	-3,35	-0,10	2,34	0,02
		IVET ($t = 10$)	-2,72	0,78	1,62	-0,53
B	100	ET ($t = 1$)	-7,64	-3,01	10,72	4,50
		ET ($t = 10$)	-5,98	-0,98	7,21	0,74
		IVET ($t = 1$)	-5,77	-2,31	4,53	-0,10
		IVET ($t = 10$)	-5,44	-1,86	5,17	-0,51
	200	ET ($t = 1$)	-2,41	-1,01	5,76	2,53
		ET ($t = 10$)	-1,29	0,18	4,30	1,91
		IVET ($t = 1$)	-1,39	-0,35	2,09	1,04
		IVET ($t = 10$)	-1,15	-0,06	2,04	0,99

EAS, échantillonnage aléatoire simple ; PPT, échantillonnage avec probabilité proportionnelle à la taille ; ET, inclinaison exponentielle (*exponential tilting*) ; IVET, inclinaison exponentielle avec variable instrumentale (*instrumental-variable exponential tilting*).

Si (37) possède une solution, celle-ci peut être exprimée comme la limite de la forme

$$w_{i(t)} \propto \prod_{s=0}^{t-1} \exp\{-\hat{U}'_{(s)}\hat{\Sigma}^{-1}_{aa(s)}U(\mathbf{x}_i)\} \qquad (38)$$

où $\hat{U}_{(s)} = \sum_{i\in A} w_{i(s)}U(\mathbf{x}_i)$, $\hat{\Sigma}_{aa(t)} = \sum_{i\in A} w_{i(t)}\{U(\mathbf{x}_i) - \bar{U}_{(t)}\}^{\otimes 2}$, $\bar{U}_{(t)} = \sum_{i\in A} w_{i(t)}U(\mathbf{x}_i)/\sum_{i\in A} w_{i(t)}$ avec le poids initial $w_{i(0)} = d_i(\hat{N}/\hat{N}_d)$. Si la solution de la condition (37) n'existe pas, nous pouvons encore utiliser les poids donnés par (38), mais l'égalité doit être relâchée. À la place, l'inégalité approximative sera satisfaite dans (37), en ce sens que $\sum_{i\in A} w_{i(t)}U(\mathbf{x}_i)$ converge vers zéro beaucoup plus rapidement que $\sum_{i\in A} w_{i(0)}U(\mathbf{x}_i)$ pour $t \geq 1$. L'égalité approximative dans (37) est appelée condition de calage approximatif.

Les estimateurs $\hat{Y}_{(t)} = \sum_{i\in A} w_{i(t)}y_i$ dans lesquels sont utilisés les poids ET en t étapes donnés par (38), y compris l'estimateur en une étape $\hat{Y}_{(1)}$, sont asymptotiquement équivalents à l'estimateur par la régression de la forme

$$\hat{Y}_{\text{reg}} = \hat{Y}_{(0)} - \hat{U}'_{(0)}\hat{\Sigma}^{-1}_{aa(0)}\hat{\Sigma}_{ay(0)},$$

où $\hat{Y}_{(0)} = \sum_{i\in A} w_{i(0)} y_i$ et $\hat{\Sigma}_{ay(0)} = \sum_{i\in A} w_{i(0)} \{U(\mathbf{x}_i) - \bar{U}_{(0)}\} y_i$. Contrairement à l'estimateur par la régression, les poids de la méthode proposée sont toujours non négatifs. En outre, en utilisant la technique de la variable instrumentale décrite à la section 3, les poids possèdent une borne supérieure. Le choix approprié de la variable instrumentale accroît également l'efficacité de l'estimateur par calage résultant.

La méthode de calage par inclinaison exponentielle est asymptotiquement équivalente à la méthode de calage par la vraisemblance empirique, mais elle est plus intéressante du point de vue des calculs, en ce sens que les dérivées partielles ne sont pas requises dans le calcul itératif. Comme les calculs sont simples, la variance de l'estimateur proposé peut être estimée facilement en utilisant une méthode de rééchantillonnage, comme celle décrite à la section 4. Une étude plus approfondie de cette approche, y compris l'estimation par intervalle, pourrait être le sujet de futurs travaux de recherche.

Remerciements

L'auteur remercie Minsun Kim de son soutien pour les calculs, ainsi que les deux examinateurs anonymes et le rédacteur associé de leurs commentaires très utiles qui lui ont permis d'améliorer considérablement la qualité de l'article. La présente étude a été financée en partie par l'entente de coopération NCRS 68-3A75-4-122 conclue entre le Natural Resources Conservation Service du US Department of Agriculture et la Iowa State University. Toutes les opinions, constatations et conclusions ou recommandations exprimées dans le présent article sont celles de l'auteur et ne reflètent pas forcément celles du USDA Natural Resources Conservation Service.

Annexe

A. Hypothèses et preuve du théorème 1

Pour commencer, supposons qu'existent les conditions de régularité suivantes :

[A-1] La densité $f(\mathbf{x};\boldsymbol{\eta})$ est deux fois dérivable par rapport à $\boldsymbol{\eta}$ pour chaque $\mathbf{x}$ et satisfait

$$\left|\frac{\partial^2 f(\mathbf{x};\boldsymbol{\eta})}{\partial\eta_i\partial\eta_j'}\right| \le K(\mathbf{x})$$

pour la fonction $K(\mathbf{x})$, telle que $E\{K(\mathbf{x})\} < \infty$, dans un voisinage de $\boldsymbol{\eta}_{0,N}$.

[A-2] L'estimateur du pseudo-maximum de vraisemblance $\hat{\boldsymbol{\eta}}$ satisfait $\sqrt{n}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_{0,N}) = O_p(1)$.

[A-3] La matrice $E\{\mathbf{s}(\boldsymbol{\eta}_{0,N})^{\otimes 2}\}$ existe et est non singulière, où $\mathbf{s}(\boldsymbol{\eta}_{0,N}) = \partial \ln f(\mathbf{x}_i;\boldsymbol{\eta})/\partial\boldsymbol{\eta}|_{\eta=\eta_{0,N}}$.

Pour prouver le théorème 1, écrivons

$$g_i(\boldsymbol{\eta}) = \frac{f(\mathbf{x}_i;\boldsymbol{\eta}_{0,N})}{f(\mathbf{x}_i;\boldsymbol{\eta})},$$

et $w_i(\boldsymbol{\eta}) = d_i g_i(\boldsymbol{\eta})$. Le poids d'échantillonnage préférentiel estimé donné par (8) peut s'écrire $w_i = w_i(\hat{\boldsymbol{\eta}})$. Un développement en série de Taylor de $N^{-1}\sum_{i\in A} d_i \mathbf{s}_i(\hat{\boldsymbol{\eta}}) = 0$ autour de $\boldsymbol{\eta}_{0,N}$ mène à

$$\mathbf{0} = \frac{1}{N}\sum_{i\in A} d_i \mathbf{s}_i(\boldsymbol{\eta}_{0,N}) + \left\{\frac{\partial}{\partial\boldsymbol{\eta}'}\frac{1}{N}\sum_{i\in A} d_i \mathbf{s}_i(\boldsymbol{\eta}_{0,N})\right\}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_{0,N}) + o_p(|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_{0,N}|).$$

Notons que le premier terme du deuxième membre de

$$\frac{1}{N}\frac{\partial}{\partial\boldsymbol{\eta}}\sum_{i\in A} d_i \mathbf{s}_i(\boldsymbol{\eta}) = \frac{1}{N}\sum_{i\in A} d_i \frac{\partial^2 f(\mathbf{x}_i;\boldsymbol{\eta})/\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}'}{f(\mathbf{x}_i;\boldsymbol{\eta})} - \frac{1}{N}\sum_{i\in A} d_i \left\{\frac{\partial f(\mathbf{x}_i;\boldsymbol{\eta})/\partial\boldsymbol{\eta}}{f(\mathbf{x}_i;\boldsymbol{\eta})}\right\}^{\otimes 2}. \quad \text{(A1)}$$

converge vers $\int\{\partial^2 f(\mathbf{x};\boldsymbol{\eta})/\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}'\}d\mathbf{x}$ qui est égal à zéro en vertu du théorème de convergence dominée avec [A1]. Le deuxième terme converge vers $E\{\mathbf{s}(\boldsymbol{\eta}_{0,N})^{\otimes 2}\}$. Donc, en vertu de [A-2],

$$\bar{\mathbf{S}}_{0d} \equiv \frac{1}{N}\sum_{i\in A} d_i \mathbf{s}_i(\boldsymbol{\eta}_{0,N}) = O_p(n^{-1/2}) \quad \text{(A2)}$$

et

$$\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_{0N} = \hat{\boldsymbol{\Sigma}}_{ss}^{-1}\bar{\mathbf{S}}_{0d} + o_p(n^{-1/2}). \quad \text{(A3)}$$

Or, un développement en série de Taylor de $N^{-1}\hat{Y}_w = N^{-1}\sum_{i\in A} w_i(\hat{\boldsymbol{\eta}}) y_i$ autour de $\boldsymbol{\eta} = \boldsymbol{\eta}_{0,N}$ mène à

$$\frac{\hat{Y}_w}{N} = \frac{\hat{Y}_d}{N} + \left\{\frac{\partial}{\partial\boldsymbol{\eta}}\frac{1}{N}\sum_{i\in A} w_i(\boldsymbol{\eta}_{0,N}) y_i\right\}'(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_{0,N}) + o_p(|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_{0,N}|) \quad \text{(A4)}$$

en vertu de la continuité uniforme de $\partial\{\sum_{i\in A} w_i(\boldsymbol{\eta}) y_i\}/\partial\boldsymbol{\eta}$ autour de $\boldsymbol{\eta}_{0,N}$. Maintenant, en utilisant

$$\frac{\partial}{\partial\boldsymbol{\eta}} g_i(\boldsymbol{\eta}) = -\frac{f(\mathbf{x}_i;\boldsymbol{\eta})}{f(\mathbf{x}_i;\boldsymbol{\eta})} \times \frac{\partial f(\mathbf{x}_i;\boldsymbol{\eta})/\partial\boldsymbol{\eta}}{f(\mathbf{x}_i;\boldsymbol{\eta})} = -g_i(\boldsymbol{\eta}) \times s_i(\boldsymbol{\eta}),$$

où $\mathbf{s}_i(\boldsymbol{\eta}) = \partial \ln f(\mathbf{x}_i;\boldsymbol{\eta})/\partial\boldsymbol{\eta}$, nous avons

$$\frac{\partial}{\partial\boldsymbol{\eta}}\sum_{i\in A} w_i(\boldsymbol{\eta}) y_i = -\sum_{i\in A} w_i(\boldsymbol{\eta}) \mathbf{s}_i(\boldsymbol{\eta}) y_i.$$

En utilisant $w_i(\boldsymbol{\eta}_{0,N}) = d_i$ et en écrivant $\mathbf{s}_i(\boldsymbol{\eta}_{0,N}) = \mathbf{s}_{i0}$, nous avons, en vertu de (A2),

$$\frac{\partial}{\partial \boldsymbol{\eta}} \frac{1}{N} \sum_{i \in A} w_i(\boldsymbol{\eta}_{0,N}) y_i = -\frac{1}{N} \sum_{i \in A} d_i \mathbf{s}_{i0} y_i$$

$$= -\hat{\boldsymbol{\Sigma}}_{sy} + O_p(n^{-1/2}). \qquad \text{(A5)}$$

En utilisant (A5) et (A3) dans (A4), nous obtenons le résultat (9).

B. Preuve du théorème 2

Écrivons

$$\hat{\theta}(\boldsymbol{\lambda}_1) = \frac{\sum_{i \in A} d_i m_i(\boldsymbol{\lambda}_1) y_i}{\sum_{i \in A} d_i m_i(\boldsymbol{\lambda}_1)},$$

où $m_i(\boldsymbol{\lambda}_1) = \exp(\boldsymbol{\lambda}_1' \mathbf{x}_i)$. Notons que $\hat{Y}_{\mathrm{ET}(t)} = \hat{N} \hat{\theta}(\hat{\boldsymbol{\lambda}}_{1(t)})$ et que $\hat{\boldsymbol{\lambda}}_{1(t)}$ est défini dans (19). Par un développement en série de Taylor de $\hat{\theta}(\hat{\boldsymbol{\lambda}}_{1(t)}) = \hat{N}^{-1} \hat{Y}_{\mathrm{ET}(t)}$ autour de $\boldsymbol{\lambda}_1 = \mathbf{0}$ et en vertu de la continuité des dérivées partielles de $\hat{\theta}(\boldsymbol{\lambda}_1)$, nous avons

$$\hat{\theta}(\hat{\boldsymbol{\lambda}}_{1(t)}) = \hat{\theta}(\mathbf{0}) + \dot{\theta}(\mathbf{0})'(\hat{\boldsymbol{\lambda}}_{1(t)} - \mathbf{0}) + o_p(|\hat{\boldsymbol{\lambda}}_{1(t)} - \mathbf{0}|), \qquad \text{(B1)}$$

où $\dot{\theta}(\boldsymbol{\lambda}) = \partial \hat{\theta}(\boldsymbol{\lambda}) / \partial \boldsymbol{\lambda}$. Comme la convergence de $\hat{\boldsymbol{\lambda}}_{1(t)}$ est d'ordre quadratique et que l'estimateur en une étape satisfait $\hat{\boldsymbol{\lambda}}_{1(1)} = O_p(n^{-1/2})$, l'équation (22) peut s'écrire

$$\hat{\boldsymbol{\lambda}}_{1(t)} = \left\{ \hat{N}_d^{-1} \sum_{i \in A} d_i (\mathbf{x}_i - \bar{\mathbf{X}}_d)^{\otimes 2} \right\}^{-1} (\hat{N}^{-1} \mathbf{X} - \bar{\mathbf{X}}_d)$$

$$+ o_p(n^{-1/2}). \qquad \text{(B2)}$$

Notons que

$$\dot{\theta}(\boldsymbol{\lambda}_1) = \left\{ \sum_{i \in A} d_i m_i(\boldsymbol{\lambda}_1) \right\}^{-1} \sum_{i \in A} d_i \dot{m}_i(\boldsymbol{\lambda}_1) \{ y_i - \hat{\theta}(\boldsymbol{\lambda}_1) \}$$

où $\dot{m}_i(\boldsymbol{\lambda}_1) = \partial m_i(\boldsymbol{\lambda}_1) / \partial \boldsymbol{\lambda}_1$. En utilisant $m_i(\mathbf{0}) = 1$ et $\dot{m}_i(\mathbf{0}) = \mathbf{x}_i$, nous avons $\hat{\theta}(\mathbf{0}) = \hat{Y}_d / \hat{N}_d$ et

$$\dot{\theta}(\mathbf{0}) = \hat{N}_d^{-1} \sum_{i \in A} d_i (\mathbf{x}_i - \bar{\mathbf{X}}_d) y_i. \qquad \text{(B3)}$$

Par conséquent, en insérant (B2) et (B3) dans (B1), nous avons

$$\hat{\theta}(\hat{\boldsymbol{\lambda}}_{1(t)}) = \frac{\hat{Y}_d}{\hat{N}_d} + \left(\frac{\mathbf{X}}{\hat{N}} - \bar{\mathbf{X}}_d \right)' \left\{ \sum_{i \in A} d_i (\mathbf{x}_i - \bar{\mathbf{X}}_d)^{\otimes 2} \right\}^{-1} \sum_{i \in A} d_i (\mathbf{x}_i - \bar{\mathbf{X}}_d) y_i + o_p(n^{-1/2}),$$

ce qui prouve (23).

Bibliographie

Beaumont, J.-F., et Bocci, C. (2008). Another look at ridge calibration. *Metron*, LXVI, 5-20.

Breidt, F.J., Claeskens, G. et Opsomer, J.D. (2005). Model-assisted estimation for complex surveys using penalised splines. *Biometrika*, 92, 831-846.

Chambers, R.L. (1996). Robust case-weighting for multipurpose establishment surveys. *Journal of Official Statistics*, 12, 3-32.

Chen, J., et Qin, J. (1993). Empirical likelihood estimation for finite populations and the effective usage of auxiliary information. *Biometrika*, 80, 107-116.

Chen, J., et Sitter, R.R. (1999). A pseudo empirical likelihood approach to the effective use of auxiliary information in complex surveys. *Statistica Sinica*, 9, 385-406.

Chen, J., Variyath, A.M. et Abraham, B. (2008). Adjusted empirical likelihood and its properties. *Journal of Computational and Graphical Statistics*, 17, 426-443.

Deville, J.-C., et Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.

Deville, J.-C., Särndal, C.-E. et Sautory, O. (1993). Generalized raking procedure in survey sampling, *Journal of the American Statistical Association*, 88, 1013-1020.

Efron, B. (1981). Nonparametric standard errors and confidence intervals. *Canadian Journal of Statistics*, 9, 139-172.

Estevao, V.M., et Särndal, C.-E. (2000). A functional approach to calibration. *Journal of Official Statistics*, 16, 379-399.

Folsom, R.E. (1991). Exponential and logistic weight adjustment for sampling and nonresponse error reduction. Dans *Proceedings of the Section on Social Statistics*, American Statistical Association, 197-202.

Fuller, W.A. (2002). Estimation par régression appliquée à l'échantillonnage. *Techniques d'enquête*, 28, 5-25.

Fuller, W.A. (2009). *Sampling Statistics*. Hoboken, New Jersey : John Wiley & Sons, Inc.

Givens, G.H., et Hoeting, J.A. (2005). *Computational Statistics*. Hoboken, New Jersey : John Wiley & Sons, Inc.

Henmi, M., Yoshida, R. et Eguchi, S. (2007). Importance sampling via the estimated sampler. *Biometrika*, 94, 985-991.

Imbens, G.W. (2002). Generalized method of moments and empirical likelihood. *Journal of Business and Economic Statistics*, 20, 493-506.

Isaki, C., et Fuller, W.A. (1982). Survey design under the regression superpopulation model. *Journal of the American Statistical Association*, 77, 89-96.

Kim, J.K. (2009). Calibration estimation using empirical likelihood in survey sampling. *Statistica Sinica*, 19, 145-157.

Kim, J.K., et Park, M. (2010). Calibration estimation in survey sampling. *Revue Internationale de Statistique*, Sous presse.

Kott, P.S. (2003). A practical use for instrumental-variable calibration. *Journal of Official Statistics*, 19, 265-272.

Kott, P.S. (2006). Utilisation de la pondération par calage pour la correction de la non-réponse et des erreurs de couverture. *Techniques d'enquête*, 32, 149-160.

Kitamura, Y., et Stutzer, M. (1997). An information-theoretic alternative to generalized method of moments estimation. *Econometrica*, 65, 861-874.

Park, M., et Fuller, W.A. (2009). The mixed model for survey regression estimation. *Journal of Statistical Planning and Inference*, 139, 1320-1331.

Rao, J.N.K., et Singh, A. (1997). A ridge shrinkage method for range restricted weight calibration in survey sampling. Dans *Proceedings of the Section on Survey Research Methods*, American Statist Association, 57-64.

Rust, K.F., et Rao, J.N.K. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, 5, 283-310.

Särndal, C.-E. (2007). La méthode de calage dans la théorie et la pratique des enquêtes. *Techniques d'enquête*, 33, 113-135.

Särndal, C.-E., Swenson, B. et Wretman, J.H. (1992). *Model Assisted Survey Sampling*. New York : Springer.

Tillé, Y. (1998). Estimation in surveys using conditional probabilities: Simple random sampling. *Revue Internationale de Statistique*, 66, 303-322.

Wolter, K.M. (2007). *Introduction to Variance Estimation*. 2ième Éd. New York : Springer-Verlag.

Wu, C., et Rao, J.N.K. (2006). Pseudo empirical likelihood ratio confidence intervals for complex surveys. *Canadian Journal of Statistics*, 34, 359-375.