
CALAGE À POIDS BORNÉS : QUE FAIT-ON AU JUSTE ?

Jean-Claude DEVILLE

ex Crest/ENSAI

deville.jean-claude@orange.fr

Mots-clés : calage, pondérations, choix de distance, robustesse

Résumé :

L'idée de base du calage est simple et unique : retrouver dans un échantillon des proportions connues dans la population. Si, de plus, l'effectif de la population est connu, par 'règle de trois' on retrouve aussi des totaux. La démarche de [1] et [2] consiste à partir d'un échantillon probabiliste muni d'un estimateur linéaire sans biais canonique. On en modifie les poids le moins possible au sens d'une certaine distance sous contraintes de respecter les totaux connus.

Certains poids peuvent alors être négatifs, d'autres exagérément gros, d'où l'idée, appréciée par les praticiens, de contraindre les variations relatives de poids à appartenir à un intervalle $[m, M]$. On constate alors empiriquement que les valeurs limites de m et M ne dépendent pas de la distance utilisée et que les poids limite en dépendent très peu. De plus, la plupart d'entre eux valent m ou M fois les poids initiaux.

Cette présentation explique le phénomène pour une classe de calages qui comprend tous les 'calages ordinaires' de [1] et [2]. Elle ouvre aussi la porte à des généralisations potentiellement intéressantes du calage à poids bornés utilisant des bornes plus flexibles que des constantes, dépendant des valeurs des variables de calage en particulier.

Le calage généralisé ([3] et [4]), qui ne repose pas sur une minimisation de distance, peut faire apparaître une difficulté gênante : l'existence de plusieurs solutions distinctes aux équations de calage et de poids 'interdits'. On apporte quelques éléments de réflexion sur ce problème pour lequel on (je) manque d'expérience pratique et d'approche théorique.

Bibliographie

- [1] Deville, J-C., Särndal, C. E., « Calibration estimators in survey sampling ». *Journal of the American statistical Association* 87, 376–382, (1992).
- [2] Deville, J-C., Särndal, C. E., Sautory, O., « Generalized raking procedures in survey sampling ». *Journal of the American statistical Association* 88, 1013–1020, (1993).
- [3] Deville, J-C., « La correction de la non-réponse par calage ou par échantillonnage équilibré », *Exposé présenté au Congrès de l'ACFAS, Sherbrooke, Quebec, (1998)*.
- [4] Deville, J-C., « La correction de la non-réponse par calage généralisé », *Actes des Journées de Méthodologie Statistique de l'Insee, Paris, France, (2002)*.