

# **Théorie et application de l'estimation sur petits domaines**

**Jean-François Beaumont**

Statistique Canada  
100 Tunney's Pasture Driveway, Édifice R.H. Coats, 25<sup>ième</sup> étage  
Ottawa, Canada, K1A 0T6  
[jean-francois.beaumont@canada.ca](mailto:jean-francois.beaumont@canada.ca)

Les enquêtes menées par les Agences nationales de statistique permettent généralement de produire des estimations fiables de paramètres de population (ex. : le taux de chômage) lorsque les tailles d'échantillon dans les sous-groupes d'intérêt de la population, appelés les domaines (ex. : des villes), sont suffisamment grandes. Les utilisateurs modernes de données d'enquête sont de plus en plus friands d'informations à des niveaux extrêmement fins (ex. : le croisement de villes et de catégories de profession). Les méthodes d'estimation traditionnelles ne peuvent habituellement pas combler ces besoins sans augmenter drastiquement la taille de l'échantillon et ainsi les coûts de collecte. Une alternative à l'augmentation de la taille de l'échantillon est l'utilisation de méthodes d'estimation sur petits domaines. Ces méthodes peuvent être utilisées pour produire des estimations fiables pour des domaines très fins sans coûts supplémentaires. Elles reposent toutefois sur des hypothèses de modèle. La validation de ces hypothèses est une étape critique dans la production d'estimation sur petits domaines.

L'atelier est divisé en quatre parties principales. Dans la première partie, on décrit le problème et quelques concepts clés tels que la différence entre un estimateur direct et indirect. On tente de démystifier l'apparente « magie » des méthodes d'estimation sur petits domaines et les risques associés à leur utilisation. Dans la deuxième partie, la plus importante de l'atelier, on décrit le modèle de Fay-Herriot qui est un modèle au niveau des domaines très fréquemment utilisé en pratique. On décrit aussi comment est construit le meilleur prédicteur linéaire sans biais. On présente plusieurs diagnostics permettant de valider le modèle de Fay-Herriot et on illustre l'utilisation de ces diagnostics au moyen de l'Enquête canadienne sur la population active. Dans la troisième partie, on décrit brièvement l'approche bayésienne hiérarchique pour le modèle de Fay-Herriot. Finalement, dans la quatrième partie, on présente le meilleur prédicteur linéaire sans biais dans le contexte d'un modèle au niveau des unités.

L'objectif du cours n'est pas de présenter en détail tous les modèles et méthodes d'estimation sur petits domaines qu'on retrouve dans la littérature. On met l'emphase sur les modèles et méthodes les plus fréquemment utilisés et on porte une attention particulière sur les hypothèses qu'ils sous-tendent et leur validation.

## **Note biographique**

Jean-François Beaumont travaille à Statistique Canada depuis 1996. Il est présentement un Chef dans une Section sur la recherche et l'innovation en statistique où il fournit des services de consultation pour les programmes de statistiques sociales et économiques. Jean-François est aussi impliqué dans quelques projets de recherche appliquée. Sa recherche porte sur différents sujets dont les principaux sont : l'estimation sur petits domaines, l'intégration de données, l'estimation robuste, la méthode du bootstrap, le traitement des valeurs manquantes et les problèmes d'estimation et d'inférence en général.