

# TRAITEMENT DES VALEURS INFLUENTES DANS LES ENQUÊTES

Jean-François Beaumont <sup>1</sup> & David Haziza <sup>2,3</sup>

<sup>1</sup> *Division de la recherche et de l'innovation en statistique, Statistique Canada, 100 Tunney's Pasture Driveway, R.H. Coats Bldg., 16-th floor, Ottawa, Ontario, K1A 0T6, Canada, jean-francois.beaumont@statcan.gc.ca*

<sup>2</sup> *Département de mathématiques et de statistique, Université de Montréal, CP 6128, succ. Centre-ville, Montréal, Québec, H3C 3J7, Canada, haziza@dms.umontreal.ca*

<sup>3</sup> *Crest (Ensaï), rue Blaise Pascal, 35172 BRUZ*

Lorsque la distribution des variables d'une enquête est fortement asymétrique, il est probable que le méthodologue soit confronté à la présence de valeurs influentes. Ces dernières sont problématiques car elles conduisent généralement à des estimateurs très instables. Le problème est particulièrement criant dans les enquêtes auprès des entreprises qui recueillent des variables économiques dont les distributions sont fortement asymétriques. Au cours de l'atelier, nous tenterons de répondre aux questions suivantes :

- Qu'est-ce qu'une valeur influente dans le contexte des enquêtes ?
- Comment quantifier l'influence d'une unité ?
- Comment réduire l'impact des unités ayant une grande influence à l'étape de l'estimation ?

À la fin de l'atelier, les participants devraient être en mesure de mieux comprendre :

- les différences entre certains concepts reliés à l'estimation robuste dont les données aberrantes (outliers), extrêmes et influentes sous différentes approches d'inférence ;
- les enjeux méthodologiques liés aux données aberrantes, extrêmes et influentes ;
- comment traiter ces types de données.

Le cours sera divisé en huit parties :

1. Statistique classique
2. L'univers des enquêtes
3. Approche basée sur un modèle
4. Approche basée sur le plan de sondage
5. Influence et biais conditionnel
6. Traitements usuels
7. Estimation robuste avec le biais conditionnel
8. Considérations pratiques

## Notes biographiques

Jean-François Beaumont travaille à Statistique Canada depuis 1996. Il est présentement un Chef dans la Division de la recherche et de l'innovation en statistique où il fournit des services de consultation autant pour les enquêtes sociales que pour celles auprès des entreprises. Jean-François est aussi impliqué dans quelques projets de recherche appliquée. Sa recherche porte sur différents sujets dont les principaux sont : l'estimation robuste, l'imputation, la non-réponse, la méthode du bootstrap et les problèmes d'estimation et d'inférence en général.

David Haziza est professeur agrégé au département de mathématiques et de statistique de l'Université de Montréal et actuellement professeur invité au CREST/ENSAI. Il passe une journée par semaine à Statistique Canada afin de faire de la consultation, de l'enseignement et de la recherche. Ses principaux thèmes de recherche sont : l'estimation robuste, l'imputation, la non-réponse, l'estimation de la variance.