

---

## MÉTHODES D'ESTIMATION SUR PETITS DOMAINES POUR L'INDICATEUR AROPE (AT-RISK OF POVERTY OR SOCIAL EXCLUSION) AU NIVEAU RÉGIONAL À PARTIR DE L'ENQUÊTE SRCV

Olivier SAUTORY (\*)

(\*) Insee, Direction de la méthodologie et de la coordination statistique et internationale

[olivier.sautory@insee.fr](mailto:olivier.sautory@insee.fr)

**Mots-clés** : Petits domaines, calage, indicateurs de pauvreté

---

### Résumé

Le futur règlement européen sur les statistiques sociales (règlement Integrated European Social Statistics ou IESS), en cours de discussion au Conseil européen, va introduire des contraintes en matière de précision pour 12 indicateurs, deux d'entre eux étant définis au niveau régional (dans la nomenclature « NUTS 2 », qui correspond au périmètre des anciennes régions) : l'estimation trimestrielle du nombre de chômeurs, à partir de l'enquête emploi en continu, et l'indicateur de pauvreté AROPE (at-risk of poverty or social exclusion), à partir du dispositif Statistique sur les revenus et les conditions de vie (SRCV). Concernant cet indicateur, il faudrait augmenter significativement la taille de l'échantillon actuel pour respecter les contraintes de précision. C'est pourquoi l'utilisation de méthodes d'estimation sur petits domaines est privilégiée.

P. Ardilly ([2]) a proposé une méthode d'estimation, dite « synthétique » dans la théorie de l'estimation sur petits domaines, qui n'est pas fondée sur une modélisation explicite : elle utilise une technique de calage (équivalente à la méthode d'estimation par régression généralisée) de l'échantillon national sur les structures régionales associées à des variables auxiliaires bien corrélées aux différentes formes de pauvreté, et présentes dans des sources auxiliaires exhaustives (recensement de la population, fichier des revenus disponibles localisés (RDL), bénéficiaires de l'allocation de solidarité aux personnes âgées (ASPA)).

La communication reviendra sur l'utilisation de cette méthode, qui a l'avantage de produire des jeux de poids – ici régionaux - indépendants des variables d'intérêt (ce que réclame Eurostat), en particulier sur les points suivants :

- serait-il légitime d'utiliser des méthodes de calage alternatives à celle proposée, qui assurent des poids toujours positifs ?
- quelles techniques peut-on mettre en œuvre pour « apprécier » le biais potentiel inhérent à ces méthodes qui reposent sur des hypothèses difficiles à vérifier : mesure de la qualité du modèle, graphiques, validation par une source externe... ?
- peut-on estimer de façon fiable la précision (ici l'erreur quadratique moyenne) des indicateurs régionaux ?

Par ailleurs, une comparaison sera faite avec des méthodes d'estimation sur petits domaines reposant sur des modélisations explicites « classiques » : au niveau régional (modèle de « Fay-Herriot »), au niveau individuel (Battese, Harter, Fuller).

## **Bibliographie**

- [1] Ardilly, P. (2006), *Panorama des principales méthodes d'estimation sur les petits domaines*, Documents de travail Insee N°M0602.
- [2] Ardilly, P. (2015), *Regional estimates of poverty indicators based on a calibration technique*, Statistical working papers, Eurostat, 2015
- [3] Battese, G. E., Harter, R. M. and Fuller, W. A. (1988). An error components model for prediction of county crop area using survey and satellite data. *Journal of the American Statistical Association*, 83, 28–36.
- [4] Fay, R. E. and Herriot, R. A. (1979). Estimation of Income for Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association*, 74, 269-277.
- [5] Rao, J.N.K. and Molina, I. (2013), *Small area estimation*, Wiley.