

ESTIMATION DE LA VARIANCE LIÉE AU NOUVEAU PLAN DE SONDAGE DE L'ENQUÊTE EMPLOI EN CONTINU (EEC)

Lionel Delta ¹ & Nicolas Paliod ¹

¹ *Insee, Division Sondages*

Résumé. L'Institut National de la Statistique et des Études Économiques (Insee) a réformé à partir de 2019 et pour la décennie à suivre, le plan de sondage de l'Enquête Emploi en continu (EEC) en France métropolitaine.

Une des particularités de l'EEC est qu'il s'agit d'une enquête aréolaire, c'est-à-dire que son échantillon est un ensemble de zones géographiques compactes appelées grappes, regroupées en secteurs comprenant six ou sept grappes. Le tirage des secteurs se fait au sein d'unités de coordination qui sont tirées lors d'une première phase de tirage ; le tirage des secteurs, puis celui des grappes constituent alors les phases suivantes.

Comme chacune de ces phases ajoute de l'aléa dans les estimations, le calcul de variance dans l'enquête Emploi nécessite de savoir exprimer la variance générée par chaque phase de tirage puis de savoir articuler l'ensemble de ces contributions pour en déduire la variance totale d'estimation.

Par rapport au plan de sondage précédent, le recours à des unités de coordination est complètement nouveau conduisant à un tirage des secteurs, non plus en une seule étape mais en plusieurs phases. En outre, le tirage des unités de coordination se fait de façon indirecte à partir de l'échantillon-maître - l'échantillon de premier degré - des autres enquêtes ménages. De ce fait il ne s'agit pas d'un plan de sondage classique auquel on peut spécifiquement associer des formules d'estimation de variance décrites dans la littérature.

Des études méthodologiques ont permis d'entériner le principe d'estimation de la variance de tirage des secteurs selon une formule de décomposition de la variance valable pour les plans de sondage à plusieurs degrés avec une utilisation particulière de l'estimateur de Sen-Yates-Grundy en première phase pour prendre en compte le tirage unités de coordination.

Mots-clés. Estimation de variance, échantillon équilibré, sondage indirect, échantillonnage à deux degrés.

Bibliographie

Breidt, F.J. and Chauvet (2011), Improved variance estimation for balanced samples drawn via the cube method, *Journal of statistical planning and inference*, No. 1414, pp. 479-487.

Caron, N., Deville, J-C. et Sautory, O. (1998), *Estimation de précision de données issues d'enquêtes : document méthodologique sur le logiciel POULPE*, Document de travail de l'Insee, Série Méthodologie statistique, No 9806.

- Chauvet, G. (2009), Stratified balanced sampling, *Survey Methodology*, Vol. 35, No. 1, pp. 115-119.
- Costa, L., Merly-Alpa, T. et Chevalier, M. (2018), Le renouvellement de l'échantillon Emploi : améliorations et évolutions, *Actes des Journées de Méthodologie Statistique de 2018*, Insee.
- Deville, J-C. et Lavallée, P. (2006), Sondage indirect : les fondements de la méthode généralisée du partage des poids, *Technique d'enquêtes*, Vol. 32, No 2, p. 185.
- Deville, J-C. and Tillé, Y. (2004), Efficient Balanced Sampling : The Cube Method, *Biometrika*, Vol 91, No 4, pp. 893-912.
- Deville, J-C. and Tillé, Y. (2005), Variance approximation under balanced sampling, *Journal of Statistical Planning and Inference*, Vol 128, pp. 569-591.
- Grafström, A. and Tillé, Y. (2013), Doubly balanced spatial sampling with spreading and restitution of auxiliary totals, *Environmetrics 14*, pp. 120-131.
- Paliot, N., Chevalier, M. et Deroyon, T. (2018), Coordination spatiale d'échantillons : applications à l'EEC et l'Echantillon-Maître, *Actes des Journées de Méthodologie Statistique de 2018*, Insee.
- Rao, J.N.K. (1975), Unbiased variance estimation for multi-stage designs, *Sankhya, C*, 37, pp. 133-139.
- Yates, F. and Grundy, P.M. (1953), Selection without replacement from within strata with probability proportional to size, *Journal of the Royal Statistical Society, Ser. B*, 15, pp. 253-261.