

Autour de la coordination d'échantillons poissonniens

Lionel QUALITÉ¹

L'Office Fédéral de la Statistique (Suisse), utilise un système de sélection d'enquêtes coordonnées développé par Qualité (2009). Il s'agit d'une extension de la méthode de Brewer et al. (1972), au cas de plus de deux enquêtes. Chaque échantillon transversal sélectionné avec ce système provient d'un plan de Poisson.

La variabilité de la taille des échantillons qui en découle nécessite de nouvelles procédures d'allocation des tailles d'échantillons visées dans des domaines pour remplacer celles jusque là utilisées pour des plans stratifiés. La précision moyenne d'un estimateur calé reste proche de celle obtenue avec un plan de taille fixe. Ce n'est donc pas la raison pour laquelle il faudrait modifier les procédures d'allocation. Par contre, un problème particulier introduit avec ces plans de Poisson, et qui n'était pas présent avec les plans stratifiés de taille fixe, est le risque de tirer un échantillon dont la taille est bien en dessous de ce qui avait été visé dans certains domaines.

Ce risque est en fait toujours présent lorsque la non-réponse est considérée comme une deuxième phase de sondage par un plan bernoullien dans des domaines. On peut tenter de le limiter en contrôlant, via les probabilités d'inclusion, la probabilité d'obtenir un échantillon trop petit dans les domaines. Le prix à payer est l'éloignement avec une allocation optimale que cela entraîne. Selon la situation et la manière dont l'on s'y prend, on constate que ce coût est faible comparativement à celui subi si l'on veut contrôler le risque lié à la non-réponse.

Un autre problème pour lequel nous présentons notre solution actuelle est l'adaptation du système aux enquêtes pour lesquelles une dépendance dans la sélection des unités est exigée (par exemple pour les enquêtes où l'on ne veut pas sélectionner plus d'une unité par ménage). Ces enquêtes ne peuvent être directement implémentées avec notre système. Nous montrons comment nous utilisons tout de même le système de coordination comme une des phases de tirage dans un plan en plusieurs phases qui satisfait les contraintes voulues.

Malgré la simplicité du plan de Poisson, le calcul des probabilités d'inclusion conditionnelles à chaque phase pour obtenir les probabilités finales désirées est en général non trivial lorsque les unités d'un même ménage reçoivent des probabilités d'inclusion inégales. Jusqu'à maintenant, nous avons tout de même pu fournir des solutions acceptables pour les différentes enquêtes qui ont été tirées, tout en conservant une très bonne coordination entre les enquêtes.

En termes d'impact, l'utilisation de tirages coordonnés a eu des effets modestes dans les enquêtes auprès des entreprises, du fait de la configuration de ces enquêtes : strates exhaustives ou au contraire probabilités d'inclusion très basses. Dans les enquêtes auprès des personnes, par contre, l'effet est très sensible puisqu'après 6 mois de fonctionnement, il permet d'éviter que 7% environ des unités sélectionnées pour de nouveaux échantillons soient des unités qui ont déjà été sollicitées en 2011.

¹ Office Fédéral de la Statistique et Université de Neuchâtel, Suisse

Références

Brewer, K., Early, L., et Joyce, S. (1972). *Selecting several samples from a single population*. Australian Journal of Statistics, 3:231--239.

Qualité, L. (2009). Unequal probability sampling and repeated surveys. Thèse de Doctorat. Université de Neuchâtel, Suisse.