

# Pondération et imputation dans une enquête longitudinale. Comparaison des méthodes et applications dans le cadre de la cohorte Elfe

## Section : Méthodologie

### Code d'identification : 1071

#### Porteur du projet, laboratoire de rattachement

Thierry Siméon, Ined, Aubervilliers

#### Résumé

Lorsqu'on réalise un sondage, afin d'inférer les données recueillies à l'ensemble de la population cible initiale, on a pour habitude d'attribuer à chaque répondant un poids calculé en prenant en compte la méthode de tirage et les caractéristiques des non-répondants. Dans les enquêtes longitudinales, une fois l'échantillon initial sélectionné au temps  $t_0$ , il existe une attrition plus ou moins importante. Certaines personnes, répondant à la première vague d'enquête, ne répondent plus aux enquêtes suivantes ou, plus généralement, participent à certaines vagues d'enquêtes et pas à d'autres. Pour prendre en compte cette attrition, on attribue un poids aux seuls répondants à chaque temps d'enquête  $t_n$ . Ce poids, qui représente la probabilité de non-participation au temps  $t_n$ , ne peut être modélisé que grâce aux variables connues au temps  $t_0$ . En effet, puisque certains répondants n'ont pas participé à toutes les vagues, il n'est pas possible d'utiliser les données intermédiaires pour modéliser cette attrition, même s'il semble évident que certaines variables du temps  $t_{n-1}$  permettraient de mieux expliquer la non-réponse au temps  $t_n$ . Si on se limitait aux seuls répondants à toutes les vagues, on limiterait trop fortement le nombre d'individus considéré. L'objectif de l'étude est de proposer une nouvelle méthode permettant de prendre en compte des données connues à chaque temps d'enquête pour mieux tenir compte les trajectoires des répondants / non-répondants dans la modélisation de la non-réponse.