

L'homogamie des couples – mise en œuvre de diverses méthodes de traitement de la non-réponse et analyse de leurs effets sur la mesure de l'homogamie.

Mélanie VANDERSCHULDEN

Insee - DG, division enquêtes et études démographiques.

Lorsque le nombre de variables non renseignées est important pour un certain nombre d'individus, on peut assimiler les observations correspondantes à de la non-réponse totale, que l'on corrige par pondération. Pour corriger la non-réponse partielle, il existe différentes méthodes d'imputation par substitution. Le hot-deck séquentiel, facile à mettre en œuvre et efficace dans le cas de variables à imputer peu nombreuses et indépendantes, n'est pas la méthode idéale lorsqu'on doit imputer une dizaine de variables potentiellement corrélées. Le hot-deck aléatoire permet de tenir compte des éventuels liens entre variables à imputer, en utilisant un unique donneur pour imputer les différentes variables manquantes pour une même observation. La constitution de classes d'imputation réduit le caractère aléatoire du choix du donneur, mais ne permet pas de le rationaliser au maximum. Le hot-deck métrique, qui consiste à choisir le donneur le plus proche du receveur au sens d'une distance, semble la méthode la plus adaptée. On teste chacune de ces méthodes sur des données de l'enquête « étude de l'histoire familiale » de 1999, que l'on souhaite corriger de la non-réponse, en vue d'une étude portant sur le choix du conjoint. L'objectif visé est la comparaison de leurs avantages et inconvénients et de leurs effets sur les résultats de l'étude, notamment sur la mesure de l'homogamie.