

Cours de David Haziza "Traitement de la non-réponse totale et partielle dans les enquêtes" – 22 et 23 juin 2009

Traitement de la non-réponse totale et partielle dans les enquêtes

La non-réponse est inévitable dans la grande majorité des enquêtes. Les statisticiens d'enquêtes distinguent essentiellement deux types de non-réponse: la non-réponse totale et la non-réponse partielle. La non-réponse totale survient lorsque toutes les variables d'intérêt sont manquantes ou lorsque la quantité d'information utilisable est jugée insuffisante. La non-réponse partielle survient lorsque certaines variables (mais pas toutes) sont manquantes. La non-réponse totale est habituellement traitée au moyen d'une procédure d'ajustement des poids de sondage alors que l'imputation est utilisée pour compenser pour la non-réponse partielle. L'objectif premier des méthodes de traitement (pour la non-réponse totale ou partielle) est la réduction du biais de non-réponse qui survient lorsque les répondants et les non-répondants n'ont pas les mêmes caractéristiques par rapport aux variables mesurées dans l'enquête.

L'atelier comportera trois modules et inclura du temps pour des questions et des discussions après chaque module.

Dans le premier module, nous discuterons brièvement les causes et la prévention de la non-réponse. De plus, l'Estimation ponctuelle en l'absence de non-réponse sera présentée. Les concepts de biais, variance et erreur quadratique moyenne dans le cadre basés sur le plan de sondage seront expliqués. L'estimateur de Horvitz-Thompson et les estimateurs de calage (qui incluent l'estimateur post-stratifié comme cas particulier) seront exposés. Finalement, le lien entre l'échantillonnage à deux-phases et la non-réponse sera expliquée, ce qui nous amènera à discuter du concept du mécanisme de non-réponse (incluant les concepts de mécanismes ignorables et non-ignorables).

Dans le deuxième module, nous mettrons l'accent sur le traitement de la non-réponse totale. Deux types de procédures d'ajustement des poids seront présentées: (i) l'ajustement par l'inverse de la probabilité de réponse estimée et (ii) l'ajustement par calage incluant le calage généralisé. La construction des classes de re-pondération sera présentée.

Le dernier module portera d'abord sur le traitement de la non-réponse partielle au moyen de l'imputation simple : des méthodes d'imputation utilisées en pratique seront présentées et leurs propriétés seront discutées. La construction des classes d'imputation sera présentée. Les principes sous-jacents à l'estimation de la variance en présence de données imputées seront expliqués incluant les techniques de ré-échantillonnages telles que le jackknife et le bootstrap. Finalement, l'imputation multiple ainsi que ses propriétés dans le contexte des enquêtes seront présentées.