

Note sur le calcul des coefficients de pondération

Plan de sondage

Le plan de sondage retenu fait référence à un sondage aléatoire à deux degrés et à couverture nationale, qui admet les districts de recensement ruraux (DR), comme les unités primaires (UP) et les ménages agricoles comme les unités secondaires (US). La méthode consiste à répartir la population statistique (ménages agricoles) dans les unités primaires de sorte que chacune d'elles soit rattachée sans ambiguïté à une UP bien déterminée. Le tirage de l'échantillon se fait alors en deux temps.

- Premier temps (1er degré), on tire un échantillon d'unités primaires (DR). Les unités primaires ou DR sont tirées avec des probabilités inégales et avec remise (tirage PIAR). La probabilité de sortie d'un DR à chaque tirage est choisie proportionnelle à sa taille exprimée en termes d'effectif des ménages agricoles.
- Deuxième temps (2e degré), on sélectionne au niveau de chaque unité primaire (DR) tiré au 1er degré, un échantillon d'unités secondaires. Au deuxième degré, les unités secondaires (ou ménages agricoles), sont tirées avec des probabilités égales et sans remise (tirage PESR). Un nombre constant d'unités secondaires est sélectionné dans chaque DR de l'échantillon du premier degré. Ce nombre constant d'unités secondaires est choisi égal à 5.

Calcul des poids théoriques

Pour avoir les coefficients de pondération, on prend l'inverse des probabilités d'inclusion des ménages obtenues à l'issue des deux tirages.

$$w_{ij} = \frac{1}{P_{ij}}$$

$$P_{ij} = p_i * p_j$$

Avec p_i la probabilité de tirage des DR au premier degré et p_j la probabilité de tirage des ménages au deuxième degré

Pour un sondage à probabilités inégale proportionnellement à la taille, la probabilité d'inclusion p_i est obtenu par la formule suivante :

$$p_{ik} = n_k * \frac{m_{ik}}{M_k}$$

m_{ik} est le nombre de ménage dans le DR i du département k .

M_k est le nombre total de ménages dans le département k .

n_k est le nombre de DR à tirer dans chaque département k .

Au deuxième degré, les ménages sont tirés sans remise à probabilités égales, soit 5 ménages dans chaque grappe. Sous l'hypothèse qu'au deuxième degré, les ménages sont tirés simultanément, la probabilité de tirage de 5 ménages est obtenu par la formule suivante :

$$p_i = \frac{5}{m'_{ik}}$$

m'_{ik} est le nombre de ménage dans le DR i du département k au moment de l'enquête. En principe, ce paramètre est différent du m_{ik} utilisé dans le calcul de la première probabilité d'inclusion.

Sous l'hypothèse que le nombre de ménage par DR est resté constant depuis le RGPHAE, les probabilités d'inclusion des deux tirages deviennent :

$$p_{ij} = n_k * \frac{m_{ik}}{M_k} * \frac{5}{m_{ik}} = n_k * \frac{5}{M_k}$$

La probabilité d'inclusion devient

$$p_{ij} = \frac{5 * n_k}{M_k} = \frac{\text{Nombre théorique de ménages à enquêter}}{\text{Nombre total de ménages dans le département } k}$$

Ainsi les coefficients de pondération théoriques sont obtenus par :

$$w_{ij} = \frac{\text{Nombre total de ménages dans le département } k}{\text{Nombre théorique de ménages à enquêter}}$$

Le poids doit être redressé en tenant compte des taux de réponse. Les poids corrigés sont :

$$W'_{ij} = \frac{w_{ij}}{\text{taux de réponse}}$$

Pour extrapoler les résultats issus de l'enquête, nous avons utilisé les coefficients de corrélation ci-dessus, qui est le rapport entre le nombre de ménages agricoles dans le département k divisé par le nombre de ménages effectivement enquêtés.