

M2 Statistique et Econométrie

Travaux pratiques du cours Techniques de sondage

TP 1

Estimation du nombre total de logements vacants en Haute-Garonne

Présentation des données et objectif

La population d'étude est l'ensemble des 554 communes de Haute-Garonne de moins de 10000 habitants au recensement 1999. Nous disposons pour ces communes :

- du code de la commune (CODE_ N),
- de son nom (COMMUNE),
- du code du BVQ (Bassin de Vie Quotidienne ¹) auquel la commune appartient (BVQ_ N),
- de la population (POPSDC99),
- de la variable “nombre de logements” (LOG),
- d'une variable de stratification (STRATLOG) à 4 modalités (1 à 4), obtenue à partir de la variable nombre de logements,
- et de la variable “nombre de logements vacants” (LOGVAC) en 1999.

Le fichier *rec99htegne* correspond à la base de sondage. L'objectif est d'estimer le total de logements vacants dans la population d'étude en mettant en œuvre avec les logiciels SAS et R deux plans de sondage : un plan SI et un plan PEAR.

Questions

Utiliser les logiciels SAS et R pour tirer dans le fichier *Rec99htegne* un échantillon de 50 communes selon (a) un plan SI et (b) un plan PEAR.

1. Tirer un échantillon de taille $n = 50$.
2. Estimer le nombre total de logements vacants.
3. Estimer la variance de l'estimateur précédent.
4. Donner un intervalle de confiance à 95% associé à l'estimateur du total.
5. Tirer 1000 échantillons et dessiner un histogramme des 1000 estimations.
6. Calculer la moyenne empirique, la variance empirique et le coefficient de variation des estimations.
7. Comparer les 2 plans.

¹un “bassin de vie” (selon la définition de l'INSEE) est un ensemble de communes à l'intérieur duquel les habitants et les entreprises trouvent la majorité des commerces et services dont ils ont besoin. Les BVQ comprennent les équipements de proximité.