

## Estimation par sondage simple

### Superficie et nombre d'habitants des communes de la région Auvergne-Rhône-Alpes

## 1 Objectifs, déroulement et attendus

L'objectif est de déterminer dans quelle mesure il est possible d'estimer la superficie et le nombre d'habitants de la région Auvergne-Rhône-Alpes à partir d'un échantillonnage de ses communes. Cette SAÉ se déroulera pendant 5 séances de TP, toutes encadrées. Le contenu des séances sert d'application aux contenus de la ressource R206 - Statistique inférentielle.

Les étudiants travailleront en équipes de 2 à 4 personnes, constituées au sein de leur groupe TP. Pour le vendredi 29 mars, les équipes doivent rendre sur le serveur de l'IUT un dossier :

- nommé en capitales aux noms des membres de l'équipe, par ordre alphabétique, séparés par des espaces. Exemple : « GOLDFARB RASSEL »
- contenant :
  - un document Excel nommé comme le dossier. Il contiendra l'ensemble des travaux réalisés sous Excel demandés pendant les séances.
  - un document pdf nommé comme le dossier. Il contiendra les réponses aux questions posées pendant les séances. Taille maximale de 10 pages.

Un rendu ne respectant pas la nomenclature ci-dessus ne sera pas évalué.

## 2 Séance 1 - Les communes de la région Auvergne-Rhône-Alpes

Les informations que vous allez apporter ici doivent être précisément sourcées et datées.

1. Donner la superficie et le nombre d'habitants de la région Auvergne-Rhône-Alpes. Ces valeurs serviront de référence pour l'ensemble des séances.
2. Donner les départements de la région, ainsi que les 10 villes les plus peuplées.
3. Pour chaque département et pour la région entière, donner le nombre de communes, la superficie et le nombre d'habitants.
4. Représenter ces informations graphiquement et/ou sous forme d'une ou plusieurs dataviz.
5. Calculer, pour chaque département et pour la région entière la superficie moyenne d'une commune et le nombre d'habitants moyen d'une commune.
6. Réaliser une base de données Excel comportant toutes les communes de la région avec les informations suivantes : nom, département, superficie, nombre d'habitants.
7. À partir de cette base de données, retrouver les résultats de la question 5.
8. Réaliser une décomposition de la variance de la superficie, selon les départements. Que peut-on en déduire ?
9. Réaliser une décomposition de la variance du nombre d'habitants, selon les départements. Que peut-on en déduire ?

### 3 Séance 2 - Création des échantillons

1. Étudier la distribution de la variable superficie.
2. Étudier la distribution de la variable nombre d'habitants.
3. Que peut-on dire de la représentativité d'un échantillon aléatoire vis-à-vis de chacune de ces deux variables ?
4. Construire 50 échantillons de taille  $n = 100$  individus à l'aide des commandes ALEA.ENTRE.BORNES et RECHERCHEV :
  - les 35 premiers ne doivent pas inclure la ville de Lyon
  - les 15 suivants doivent inclure la ville de Lyon

### 4 Séance 3 - Étude de la fluctuation d'échantillonnage

On note  $X$  la variable superficie.

1. Rappeler la moyenne et la variance de  $X$ .
2. Pour chacun des 50 échantillons, calculer la moyenne échantillonnale et la variance échantillonnale.
3. Représenter la distribution expérimentale de  $\bar{X}$  et calculer sa moyenne et sa variance. Retrouve-t-on expérimentalement les résultats théoriques ?
4. Donner un intervalle dans lequel se trouvent 95% des réalisations de  $\bar{X}$ . Parmi vos échantillons, combien ont une valeur  $\bar{x}$  en dehors de cet intervalle ?
5. Calculer la moyenne expérimentale de  $S^2$ . Retrouve-t-on expérimentalement le résultat théorique ?

Réaliser la même étude avec  $Y$ , la variable nombre d'habitants.

### 5 Séance 4 - Estimations

On note  $X$  la variable superficie. On souhaite estimer la moyenne  $\mu$  de la superficie d'une commune dans la région.

1. À partir de chaque échantillon, donner une estimation ponctuelle de  $\mu$ , ainsi qu'un intervalle de confiance au risque  $\alpha = 0.05$ .
2. Confronter les résultats de vos intervalles de confiance avec la vraie valeur de  $\mu$ .

Réaliser la même étude avec  $Y$ , la variable nombre d'habitants.

### 6 Séance 5 - Rédaction des rendus

Le dossier doit être rendu sur le serveur à la fin de la séance.

N'oubliez pas de rédiger une conclusion apportant une réponse à la problématique initiale : dans quelle mesure est-il possible d'estimer la superficie et le nombre d'habitants de la région Auvergne-Rhône-Alpes à partir d'un échantillonnage de ses communes ?

Profitez du temps qu'il reste pour faire le lien entre cette SAÉ et votre portfolio.