

Melun

Session : Septembre 2018

Année d'étude : Troisième année de licence économie-gestion mention administration économique et sociale

Discipline : *Sondages techniques quantitatives*
(Unité d'Enseignements Fondamentaux 1)

Titulaire(s) du cours :
M. Antoine AUBERGER

Document(s) autorisé(s) : Aucun document n'est autorisé. L'usage d'une calculatrice est autorisé.

Questions de cours :

1) Donner trois noms d'instituts de sondages.

Pour un sondage de popularité (exemple : popularité du président de la République), quelle est la méthode utilisée pour sélectionner l'échantillon ? Donner des précisions sur la méthode, sur l'échantillon et sur la stratification.

2) L'analyse factorielle des correspondances : à présenter (objectif, principes, méthode) et proposer un exemple de données.

Exercice 1 :

On souhaite estimer la quantité d'eau moyenne (exprimée en m³) consommée annuellement par les habitants d'une ville donnée de 100 000 habitants. On sélectionne par un plan simple un échantillon de 300 habitants. Les résultats obtenus sont les suivants :

$$\sum_1^{300} x_i = 14850 \text{ et } \sum_1^{300} x_i^2 = 981720$$

1) Traduire en quelques mots l'information contenue dans la formule : $\sum_1^{300} x_i = 14850$.

2) Donner un intervalle de confiance à 95% pour la quantité d'eau moyenne consommée annuellement par les habitants de cette ville.

3) On s'intéresse maintenant à la quantité totale consommée annuellement par l'ensemble des habitants de la ville. Donner une estimation, puis un intervalle de confiance à 95% pour cette quantité totale.

Exercice 2 :

Sur un campus universitaire de 3200 étudiants, un échantillon de 120 étudiants est prélevé suivant un plan de sondage aléatoire de type PESR. Parmi ces 120 étudiants, 85 étudiants admettent qu'ils utilisent internet plus d'une heure par jour. On note p_U la proportion des 3200 étudiants qui utilisent internet plus d'une heure par jour.

- 1) Donner une estimation ponctuelle de p_U .
- 2) Déterminer un intervalle de confiance pour p_U au niveau 95% (on donnera le résultat à 10^{-3} près).

Exercice 3

On considère le caractère $Y = \text{"âge"}$ en années dans la population de 4 individus :

$U = \{\text{Paul, Bernard, Yves, Matthieu}\} = \{u_1, \dots, u_4\}$. Pour tout $i \in \{1, \dots, 4\}$, soit y_i la valeur de Y pour l'individu u_i . Les résultats, en années, sont :

y_1	y_2	y_3	y_4
17	10,5	28	20,5

- 1) Calculer la moyenne-population \bar{y}_U et l'écart-type corrigé de la population S_U .
- 2) Dans un premier temps, on prélève un échantillon de 2 individus suivant un plan de sondage aléatoire simple de type PESR.
 - a) Quel est le taux de sondage ? Combien d'échantillons peut-on former ? Expliciter les.
 - b) Pour chaque échantillon ω , calculer la moyenne de l'échantillon \bar{y}_ω .
 - c) Soit \bar{y}_W la variable aléatoire réelle égale à la moyenne de l'échantillon, l'aléatoire étant dans l'échantillon considéré.
Déterminer sa loi, puis calculer son espérance et sa variance.
 - d) Avec quelle formule peut-on calculer la variance de l'estimation de \bar{y}_U (\hat{y}_U) ? Faire le calcul.
- 3) Dans un deuxième temps, on prélève un échantillon de 2 individus suivant un plan de sondage aléatoire de type ST (stratifié) avec :
 - les 2 strates : $U_1 = \{\text{Paul}\}$ et $U_2 = \{\text{Bernard, Yves, Matthieu}\}$,
 - un individu par strate.
 - a) Combien d'échantillons peut-on former ? Expliciter les.
 - b) Pour chaque échantillon ω , calculer la moyenne de l'échantillon \bar{y}_ω .
 - c) Soit \bar{y}_W la variable aléatoire réelle égale à la moyenne de l'échantillon stratifié, l'aléatoire étant dans l'échantillon considéré.
Déterminer sa loi, puis calculer son espérance et sa variance.
 - d) Calculer la variance de \hat{y}_W à l'aide de la formule : $V(\hat{\mu}) = \frac{1}{N^2} \sum_{h=1}^H N_h \frac{N_h - n_h}{n_h} S_h^2$

Exercice 4 :

Exemple d'un tableau de distances

	a	b	c	d
a	0			
b	1	0		
c	2	1	0	
d	4	3	2	0

On utilise l'agrégation par l'ultramétrie supérieure minima :

$$D(I_1, I_2) = \max\{d(i, i'), i \in I_1, i' \in I_2\}$$

On agrège les deux parties ayant les deux points les plus proches et qui vont former une classe de diamètre maxima, les classes sont longues et fines, il y a un effet de chaînage (critère du plus proche voisin).

Construire une hiérarchie : algorithme ascendant ou agrégatif.