

Melun

Session : Janvier 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 d'intention de vote (exemple : premier tour de l'élection présidentielle française de 2017), 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 en composantes principales : à présenter (objectif, principes, méthode) et proposer un exemple de données.

Exercice 1 :

Un pépiniériste souhaite estimer la taille moyenne de ses arbustes d'une même variété. Sur les 10000 plantes de la serre, on en sélectionne 250 par sondage aléatoire simple, puis on mesure la hauteur de chacune de ces plantes. Les résultats sont les suivants (en m) :

$$\sum_1^{250} x_i = 130 \text{ et } \sum_1^{250} x_i^2 = 210$$

1) Traduire en quelques mots l'information contenue dans la formule : $\sum_1^{250} x_i = 130$.

2) Donner un intervalle de confiance à 95% pour la taille moyenne des arbustes.

Exercice 2 :

Dans une grande ville, on s'intéresse au nombre moyen de clients que peut avoir un pharmacien pendant une journée de travail. On part de l'idée a priori que plus la superficie de la pharmacie est importante, plus il a de clients. On classe donc la population des pharmacies en trois groupes : les "petites" (classe 1), les

"moyennes" (classe 2), et les "grandes" (classe 3). Par ailleurs, on suppose que l'on connaît, dans la base de sondage des pharmaciens, la classe de chacun d'entre eux. On tire par sondage aléatoire simple 5 pharmaciens dans chaque classe.

On obtient les résultats suivants :

U_h	U_1	U_2	U_3
N_h	8	15	12
n_h	5	5	5
\bar{x}_h	125	238	353
s_h^2	280,1	532,5	790,1

- 1) Comment s'appelle ce plan de sondage ?
- 2) Comment estimez-vous le nombre moyen de clients par jour et par pharmacien ?
- 3) Donner un intervalle de confiance à 95% pour le vrai nombre moyen de clients par jour et par pharmacien.
- 4) Si vous n'aviez comme contrainte que le nombre total de pharmaciens à enquêter (soit 15), procéderiez-vous comme ci-dessus ?

Exercice 3 :

Sur les 8000 employés d'une entreprise, on souhaite connaître la proportion p_U d'entre eux qui ont un ordinateur portable (seul ou avec leur famille). On décide de former 3 strates en fonction du revenu des employés. On considère alors :

- la strate U_1 : ensemble des employés à revenu faible,
- la strate U_2 : ensemble des employés à revenu modeste,
- la strate U_3 : ensemble des employés à revenu fort.

On dispose des informations suivantes :

U_h	U_1	U_2	U_3
N_h	3400	2600	2000
n_h	250	170	80
p_h	0,25	0,70	0,90

- 1) Donner une estimation ponctuelle de p_U (notée \hat{p}_U).
- 2) Donner une estimation ponctuelle de l'estimateur de p_U de la variance puis de l'écart-type de l'estimateur de p_U (notées $\text{Var}(\hat{p}_U)$ et $\hat{\sigma}_{\hat{p}_U}$).
- 3) Déterminer un intervalle de confiance pour p_U au niveau 95%.

Exercice 4 :

Exemple d'un tableau de distances

	a	b	c	d
a	0			
b	1	0		
c	4	2		
d	5	4	2	0

On utilise l'agrégation par l'ultramétrie inférieure maxima

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

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.