

Melun

**Session :** Janvier 2017

**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 quatre noms d'instituts de sondages.

Pour un sondage d'intention de vote (exemple : primaire de la droite à l'élection présidentielle française), quelle est la méthode utilisée pour sélectionner l'échantillon ? Donner des précisions sur la méthode, sur l'échantillon et sur la stratification.

2) Soit une population avec N éléments et un échantillon avec n éléments (sondage sans remise).

Soit  $S^2 = \frac{1}{N-1} \sum_1^N (x_i - \bar{X})^2$ . Que représente  $S^2$  ?

Montrer que :  $S^2 = \frac{N}{N-1} (\frac{1}{N} \sum_1^N x_i^2 - \bar{X}^2)$

Si  $S^2$  est inconnue, quelle estimation de  $S^2$  utilise-t-on pour remplacer  $S^2$  ?

3) L'analyse en composantes principales : à présenter (objectif, principes, méthode) et proposer un exemple de données.

**Exercice 1 :**

Sur un campus universitaire de 3000 étudiants, un échantillon de 150 étudiants est prélevé suivant un plan de sondage aléatoire de type PESR. Parmi ces 150 étudiants, 105 étudiants disent qu'ils ont une activité sportive plus d'une heure par semaine. On note  $p_U$  la proportion des 3000 étudiants qui ont cette activité sportive.

1) Donner une estimation ponctuelle de  $p_U$ .

2) Déterminer un intervalle de confiance pour  $p_U$  au niveau 95%.

## Exercice 2 :

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

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

$y_1$	$y_2$	$y_3$	$y_4$
21	15	16	24

- 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  $\omega$ , calculer la moyenne de l'échantillon  $\bar{y}_\omega$ .
  - 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  $\bar{y}_U$  ( $\widehat{\bar{y}}_U$ ) ? 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 :  $U_1 = \{\text{Charles, Laurent}\}$  et  $U_2 = \{\text{Bernard, Paul}\}$ ,
  - un individu par strate.
  - a) Combien d'échantillons peut-on former ? Expliciter les.
  - b) Pour chaque échantillon  $\omega$ , calculer la moyenne de l'échantillon  $\bar{y}_\omega$ .
  - 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  $\widehat{\bar{y}}_W$  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 3 :

Dans une très grande ville, on s'intéresse au nombre moyen de clients que peut avoir un kiosque à journaux pendant une journée. On part de l'idée a priori que plus un kiosque à journaux est grand, plus il a de clients. On classe donc la population des kiosques à journaux en quatre groupes : les "petits" (classe 1), les "moyens" (classe 2), les "grands" (classe 3), et les "très grands" (classe 4).

Par ailleurs, on suppose que l'on connaît, dans la base de sondage des kiosques à journaux, la classe de chacun d'entre eux. On tire par sondage aléatoire simple 10 kiosques à journaux dans chaque classe.

On obtient les résultats suivants :

	h=1	h=2	h=3	h=4
$\bar{x}_h$	23	34	45	52
$s_h^2$	8	12	14	16
$N_h$	30	50	70	25

- 1) Comment s'appelle ce plan de sondage ? Préciser un peu.
- 2) Comment estimez-vous le nombre moyen de clients des kiosques à journaux par jour et par kiosque ? (on le note  $\hat{\mu}$  et on donnera le résultat à  $10^{-2}$  près).
- 3) Donner une estimation de la variance de  $\hat{\mu}$  (on donnera le résultat à  $10^{-4}$  près).

On rappelle que dans ce cas on utilise la formule :  $V(\hat{\mu}) = \frac{1}{N^2} \sum_{h=1}^H N_h \frac{N_h - n_h}{n_h} S_h^2$

4) Donner un intervalle de confiance à 95% pour le nombre moyen de clients par jour et par kiosque à journaux (on donnera le résultat à  $10^{-2}$  près).

5) Si vous n'aviez comme contrainte que le nombre total de kiosques à journaux à enquêter (soit 40), procéderiez-vous comme ci-dessus ? Préciser les 2 méthodes possibles pour calculer le nombre de kiosques à journaux de chaque catégorie dans l'échantillon (à détailler) et faire le calcul pour une méthode.

Calculer dans ce cas une estimation de la variance de  $\hat{\mu}$ .