



Université PANTHÉON - ASSAS (PARIS II)

U.E.C.1

Droit - Economie - Sciences Sociales

4615

Session : Janvier 2018

Année d'étude : Première année de Master économie-gestion mention ingénierie économique et statistique

Discipline : **Sondages** (Unité d'Enseignements Complémentaires 1)

Titulaire(s) du cours : M. Philippe PERIE

Document(s) autorisé(s) : Formulaire

Les exercices sont indépendants et peuvent être résolus dans n'importe quel ordre.
Barème indicatif :

question de cours : 8 points (1.5,1.5,1.5,1.5,2),

exercice1 : 8 points (1.5,1.5,2,2,1)

exercice2 : 4 points (2,2)

Questions de cours :

- Quel est le principe du sondage en plusieurs degrés ?
- Quels sont ses avantages et inconvénients par rapport au sondage aléatoire simple ?
- Quel est le principe du sondage stratifié ?
- Quels sont ses avantages et inconvénients par rapport au sondage aléatoire simple ?
- Donner un exemple de mise en œuvre combinant de manière efficace les techniques de degrés et de stratification de manière à bénéficier simultanément de leurs avantages respectifs

Exercice 1 : Strates d'éléphants : Un directeur de zoo possède 100 éléphants classés en 2 catégories : mâles et femelles. (N = 60 mâles et 40 femelles) À la suite d'un sinistre, le zoo doit être fermé pendant plusieurs mois, et il faut en estimer le poids pour organiser rapidement le transport des éléphants vers d'autres zoos. Il y a 10 ans, le zoo avait fait peser tous les éléphants du troupeau (il y en avait alors 50 et cela avait pris du temps ...) et avait obtenu les résultats présentés ci-dessous

	Effectifs N_h	Moyennes \bar{Y}_h	Dispersions $S_{y_h}^2$
Mâles	30	6	4
Femelles	20	4	2.25

Poids moyen et dispersions selon les strates

Le directeur du zoo doit organiser le déménagement et n'a pas le temps de peser les 100 éléphants, il fait donc l'hypothèse que le poids moyens par sexe n'a pas changé en 10 ans. Il doit commander des convois de camions pour les transporter et doit optimiser son budget tout en limitant les risques de se tromper. On suppose que chaque camion peut transporter deux éléphants (mâles ou femelle), mais que le prix facturé du transport est au poids : 100 euros la tonne transportée.

- Calculez la dispersion dans la population de la variable 'poids de l'éléphant' il y a 10 ans
- Si le directeur procède à un tirage aléatoire de 10 éléphants, donnez la variance de l'estimateur du poids total du troupeau, et donnez en un intervalle de confiance à 95%. Donnez le niveau maximal de prix que le directeur peut s'attendre à payer avec 5% au plus de risque de se tromper.
- Même question si le directeur procède à un tirage à allocation proportionnelle
- Si le directeur procède à un tirage à allocation optimale, quels sont les effectifs dans chacune des deux strates et quelle est la variance de l'estimateur du total ? Donnez alors aussi un niveau maximal de prix à payer à 5% de risque
- Commentez

Exercice 2 : Fraude aux URSSAF, estimation du nombre de fraudeurs

L'URSSAF est chargé en France du recouvrement des cotisations de sécurité sociale et des allocations familiales. La fraude aux allocations est un sujet de manque à gagner pour l'état, et l'estimation du taux de fraude sert à suivre l'efficacité des mesures de contrôle. Lors d'enquêtes préliminaires dans lesquelles les inspecteurs épiluchent les comptes en détail, on a trouvé les valeurs suivantes, dans 3 secteurs d'activité 1,2 et 3

On note :

N_h = taille de la strate h du secteur d'activité

n_h = taille de l'échantillon dans la strate h (sondage aléatoire simple dans chaque strate)

\widehat{p}_h = estimateur de la proportion de fraudeurs dans la strate h

Taux de fraude moyen et dispersions selon les strates de secteurs d'activité

Strates	Taille du secteur d'activité N_h	Taille de l'échantillon n_h	Proportion de fraudeurs (%) \widehat{p}_h
Strate h1	70000	1000	1.5
Strate h2	20000	600	4.5
Strate h3	20000	400	5.0

- Quel est l'estimateur stratifié de la proportion globale de fraudeurs \bar{p}_{str} ?
- Donnez un estimateur sans biais de la variance $\widehat{var}(\widehat{p}_{str})$ et un intervalle de confiance à 95%