UNIVERSITE DE YAOUNDE II
THE UNIVERSITY OF YAOUNDE II

FACULTE DES SCIENCES
ECONOMIQUES ET DE GESTION
BP. 1365 YAOUNDE
CAMEROUN
www.univ-yde2.org
Tél: (237) 22 06 26 98 / Fax (237)22 23 84 28

FACULTY OF ECONOMICS AND MANAGEMENT
P.O Box 1365 YAOUNDE
CAMEROON
fseg@.univ-yde2.org

Tel: (237) 22 06 26 98/ Fax (237) 22 23 84 28

Année académique 2024-2025

Travaux dirigés de techniques de sondage et enquête

Exercice 1

On considère pour cela tous les échantillons possibles de taille 2 pris dans une population de taille $N = 5$. On connaît par ailleurs les valeurs de la variable d'intérêt Y pour chaque unité de la population, à savoir respectivement : 8, 3, 11, 4 et 7.

1. Calculer la moyenne Y et la dispersion S_Y^2 du caractère d'intérêt sur la population.
2. Lister tous les échantillons possibles de taille 2.
3. Pour chacun de ces échantillons, calculer l'estimateur $\hat{\bar{Y}}$ de la moyenne de la variable d'intérêt.
4. Vérifier que $\hat{\bar{Y}}$ estime sans biais la vraie moyenne.
5. Calculer la variance de cet estimateur $V\left(\hat{\bar{Y}}\right)$
6. Vérifier que cette variance $V\left(\hat{\bar{Y}}\right)$ coïncide avec la formule donnée par la théorie.
7. Calculer l'estimateur de variance $\hat{V}\left(\hat{\bar{Y}}\right)$ pour chacun des échantillons possibles.
8. Vérifier que $\hat{V}\left(\hat{\bar{Y}}\right)$ estime sans biais la vraie variance $V\left(\hat{\bar{Y}}\right)$.

Exercice 2. Soit la population $\{1,2,3\}$. On considère le plan de sondage suivant :
(n=2)

$$P(\{1,2\}) = \frac{1}{2} \quad P(\{1,3\}) = \frac{1}{4} \quad P(\{2,3\}) = \frac{1}{4}$$

1. Est-ce un sondage aléatoire simple ?
2. Calculer la probabilité pour que l'individu 1 fasse partie de l'échantillon. Même question pour les individus 2 et 3.
3. Calculer la valeur de l'estimateur de la moyenne pour chaque échantillon possible.
4. Vérifier que cet estimateur est biaisé.

Exercice 3

Considérons qu'on veuille estimer le total et la moyenne d'une grandeur Y dans une population U de taille N. Pour cela, on procède à un sondage aléatoire simple sans remise de taille n et on note S l'échantillon aléatoire obtenu.

1. Combien y a-t-il d'échantillons possibles ? Quelle est la probabilité de tirer chacun d'entre eux ?

2. On considère un individu k quelconque dans U. Combien y a-t-il d'échantillons contenant cet individu ? En déduire la probabilité de tirage de k.

3. On note I_k la variable aléatoire valant 1 si k appartient à l'échantillon et 0 sinon.

a. Que vaut $E(I_k)$?

b. Comment peut-on réécrire $\sum_{k\in S} Y_k$ à partir des I_k ?

4. En déduire que :

a. $\hat{t}_y = \frac{N}{n}\sum_{k\in S} Y_k$ estime sans biais le vrai total $t_y = \sum_{k\in U} Y_k$

b. et que $\hat{\bar{Y}} = \frac{1}{n}\sum_{k\epsilon S} Y_k$ estime sans biais la vraie moyenne $\bar{Y} = \frac{1}{N}\sum_{k\epsilon U} Y_k$.

5. Combien y a-t-il d'échantillons comprenant les individus identifiés k et l ? En déduire la probabilité de tirer ces deux individus conjointement. Que vaut alors $E(I_k I_l)$? En déduire $Cov(I_k, I_l)$.

6. On note $S_y^2 = \frac{1}{N-1}\sum_{k\in U}(Y_k - \bar{Y})^2$ et $f = \frac{n}{N}$. Montre que :

a. $Var(\hat{t}_y) = N(N-n)\frac{S_y^2}{n}$

b. $Var\left(\hat{\bar{Y}}\right) = (1-f)\frac{S_y^2}{n}$

7. Quel est l'intérêt du sondage sans remise par rapport au sondage avec remise ?

8. Montrer que $S^2 = \frac{1}{n-1}\sum_{k\in S}\left(Y_k - \hat{\bar{Y}}\right)^2$ estime sans biais S_y^2.

9. En déduire des estimateurs sans biais de $Var(\hat{t}_y)$ et de $Var\left(\hat{\bar{Y}}\right)$

.

Exercice 4. Les 4 pandas d'un zoo de Vienne forment une population de quatre sujets. Soit Y la variable donnant leur poids et {98.7, 102.6, 108.5, 120.3} les valeurs observées.

a. Déterminer la moyenne et la variance dans cette population
b. Combien d'échantillons de taille 3 peut-on extraire de la population ? Citez-les
c. Pour chacun de ces échantillons, calculer la moyenne et la variance
d. En supposant que les différents échantillons soient équiprobables, donner la distribution de l'estimateur de la moyenne de la population et calculer l'espérance de cette moyenne, son biais et son erreur quadratique moyenne
e. Les pandas femelles se laissent plus difficilement attraper que les mâles. La population étant constituée de trois mâles et d'une femelle, l'échantillon masculin a une probabilité qui vaut 2/5 contre 1/5 seulement pour les échantillons contenant la femelle. Donner alors la distribution de l'estimateur de la moyenne de la population et calculer l'espérance de cette moyenne, son biais et son erreur quadratique moyenne.

<u>Exercice 5</u>.

Considérons une population de quatre sujets sur lesquels on mesure une variable X dont voici les valeurs : {6.2; 9.3; 9.9; 10.7}

a. Déterminer la moyenne et la variance dans cette population ;

b. Combien d'échantillons de taille 3 peut-on extraire de la population ? Citez-les

c. Pour chacun de ces échantillons, calculer la moyenne et la variance ;

d. En supposant que les différents échantillons sont équiprobables, donner la distribution de l'estimateur de la moyenne de la population et calculer l'espérance de cette moyenne, son biais et son erreur quadratique moyenne ;

e. En supposant maintenant que l'individu n°4 soit moins disponible que les autres pour l'enquête, recalculer ces valeurs si la probabilité de l'échantillon dans lequel il n'apparaît pas est deux fois plus grande que pour les autres échantillons.

<u>Exercice 6</u>.

On veut estimer la superficie moyenne cultivée dans les fermes d'un canton rural. Sur les 2010 fermes que comprend le canton, on en tire 100 par sondage aléatoire simple. On mesure (en hectares) la surface cultivée x_k par la ferme numéro k de l'échantillon et on trouve :

$$\sum_{k=1}^{100} x_k = 2907 \qquad et \qquad \sum_{k=1}^{100} x_k^2 = 154.593$$

1. Donner la valeur de l'estimateur de la moyenne $\hat{\mu} = \bar{x}$.
2. Donner un intervalle de confiance à 95% pour μ.

<u>Exercice 7</u>.

Un pépiniériste souhaite estimer la taille moyenne des arbustes d'une même variété. Sur les 10 000 plantes de la serre, on en sélectionne 200 par sondage aléatoire simple, puis on mesure la hauteur de chacune de ces plantes. Les résultats sont les suivants (en m) :

$$\sum_{k=1}^{200} x_k = 248 \qquad et \qquad \sum_{k=1}^{200} x_k^2 = 331$$

1. Donner un intervalle de confiance à 95% pour la taille moyenne des échantillons.
2. Le pépiniériste a de bonnes raisons de penser que l'écart-type calculé sur la population de tous les arbustes se situe entre 0,25 et 0,45 m. En négligeant le taux de sondage, quelle taille d'échantillon doit-on retenir pour donner un intervalle de confiance à 95% ayant une demi-longueur d'au plus 2 cm ?