TP 6

STT2000 : Échantillonnage

À remettre le 18 novembre

1. Supposons que les valeurs de deux variables y et x sont observées pour les n unités tirées d'une population de taille N selon l'EASSR. On cherche à estimer le ratio $B = \dfrac{\bar{y}_U}{\bar{x}_U}$. Une source administrative nous permet de connaitre la vraie valeur de la variable x dans la population, $\bar{x}_U$. Quelle stratégie d'estimation utiliseriez-vous? (a) utiliser **toujours** l'estimateur $\dfrac{\bar{y}_s}{\bar{x}_U}$. (b) Utiliser **parfois** l'estimateur $\dfrac{\bar{y}_s}{\bar{x}_U}$. (c) utiliser **toujours** l'estimateur $\dfrac{\bar{y}_s}{\bar{x}_s}$. (d) utiliser **parfois** l'estimateur $\dfrac{\bar{y}_s}{\bar{x}_s}$. *Attention ! Faites une réponse complète. Considérez le biais, la variance, l'EQM. Démontrez vos résultats et discutez ce que vous obtenez.*

2. Supposons que l'on cherche à estimer le ratio de deux ratios

$$B = B_1 / B_2$$

où $B_1 = \dfrac{\bar{y}_{1U}}{\bar{x}_{1U}}$ et $B_2 = \dfrac{\bar{y}_{2U}}{\bar{x}_{2U}}$.

a) Exprimer B comme une fonction de totaux.
b) Afin d'estimer B, On tire un EASSR, s, de taille n. Un estimateur de B est défini selon

$$\hat{B} = \hat{B}_1 / \hat{B}_2$$

Utiliser une approximation par séries de Taylor du premier ordre pour montrer que $\hat{B} - B \approx \dfrac{N}{n}\sum_{i\in s} z_i$, et donner l'expression pour z_i.
c) Donner une expression de la variance approximative de $\hat{B}$.

3. **Échantillonnage aléatoire simple avec remise** (EASAR) : L'EASAR est défini comme suit : On fait n tirages indépendants dans la population U de taille N. Au tirage i, $i = 1,...,n$, on accorde la même probabilité $\dfrac{1}{N}$ d'être choisie à toute unité $i \in U$. Dans ce qui suit, supposons que l'on tire un EASAR, s, de taille n. Soit Q_i le nombre de fois que l'unité i est sélectionnée dans l'échantillon.

a) Montrer que $\pi_i = P(i \in s) = P(Q_i \geq 1) = 1 - \left(1 - \dfrac{1}{N}\right)^n$.
b) Déterminer π_{ij}. Montrer le raisonnement.
c) Quelle est la distribution de $\mathbf{Q} = (Q_1, Q_2, ..., Q_N)$? (voir Annexe A.5 dans les notes de cours)
d) On cherche à estimer le total $t_y = \sum_{i\in U} y_i$. On propose l'estimateur suivant: $\hat{t}_y = \dfrac{N}{n}\sum_{i\in U} Q_i y_i$. En utilisant le résultat trouvé en (c), montrer que $E(\hat{t}_y) = t_y$.
e) Déterminer $V(\hat{t}_y)$ en utilisant les résultats trouvés en (b). Simplifier votre expression le plus possible.