## **Exercices semaines 1 et 2\*\*\*4%**

## *Pour répondre à toutes les questions ci-dessous, vous devez utiliser Stata (et, spécifiquement, DASP, si demandé). Soyez concis(es) et clair(e)s dans vos réponses.*

## *L’examen est divisé en trois exercices (les points assignés à chaque exercice sont indiqués à côté de chaque exercice). Veuillez répondre (R) directement dans ce fichier après chaque question (Q) et veuillez joindre le fichier \*.do (do-file) que vous avez généré. Renommez ces deux fichiers en : "Exercice semaines 1-2-3 - Prénom, Nom" et veuillez les* soumettre *par la boîte de dépôt du portail de cours avant mardi le 5 février 11h59 a.m. (*[*heure du Québec*](https://www.timeanddate.com/worldclock/converter.html?iso=20190205T165900&p1=189)*).*

## **Exercice 1 (4%)\*\*\*\*4%**

Supposons que la population est composée de 12 ménages qui vivent dans les régions *A, B et C*.

|  |  |  |  |
| --- | --- | --- | --- |
| *identifier* | *region* | *income* | *hhsize* |
| 1 | A | 210 | 4 |
| 2 | A | 450 | 6 |
| 3 | A | 300 | 5 |
| 4 | A | 210 | 3 |
| 5 | B | 560 | 2 |
| 6 | B | 400 | 4 |
| 7 | C | 140 | 4 |
| 8 | C | 250 | 2 |
| 9 | C | 340 | 2 |
| 10 | C | 220 | 2 |
| 11 | C | 360 | 3 |
| 12 | C | 338 | 3 |

**Q 1.1:** À l’aide de Stata, générez le revenu par habitant (*pcinc*).

**R :** Pour générer le revenu par habitant (*pcinc*), on utilise la commande suivante :

**gen pcinc = income/hhsize**

**Q 1.2:** À l'aide de Stata, estimez le revenu moyen par habitant et le revenu total de notre population.

**R :** L’estimation du revenu moyen par habitant et le revenu total de la population sera à l’aide de la commande suivante :

**sum pcinc [aw=hhsize]**

**Q 1.3:** Supposons que le seuil de pauvreté soit égal à 100. Générez la variable « intensité de la pauvreté par habitant (*pgap*) », puis estimez sa moyenne (l’intensité de la pauvreté par habitant devrait être normalisée par le seuil de pauvreté).

**R :**

* Pour générer la variable « intensité de la pauvreté par habitant (*pgap*) », on utilise les commandes suivantes :

**gen pline = 100**

**gen pgap = 0**

**replace pgap = (pline-pcinc)/pline if (pcinc < pline)**

* Pour estimer la moyenne (l’intensité de la pauvreté par habitant devrait être normalisée par le seuil de pauvreté), la commande stata suivante sera nécessaire :

**sum pgap [aw=hhsize]**

**Q 1.4:** Refaites la question Q 1.3 avec DASP.

**R :** Pour reprendre la question Q 1.3 avec DASP, la commande suivante sera utilisée

**ifgt pcinc, pline(100) alpha(1) hsize(hhsize)**

**Q 1.5:** Supposons que le pouvoir d'achat dans la région B soit supérieur de 10% à celui de la région A et que celui de la région C soit supérieur de 30% à celui de la région A. Dans le cas où la région A est la région de référence, générez la variable (deflator) en tant qu'indice de déflation des prix, puis générez la variable de revenu réel par habitant (rpcinc).

**R :** Dans ces conditions, les commandes suivantes seront utilisées

**gen deflator = 1**

**replace deflator = 1.1 if region == 2 & deflator = 1.3 if region == 3**

**gen rpcinc = pcinc/deflator**

**Q 1.6:** Refaites les questions 1.3 et 1.4 en utilisant le revenu réel par habitant lorsque le seuil de pauvreté est de 120.

**R :**

* Pour la reprise de la 1.3, on utilisera les commandes suivantes

**sum rpcinc [aw=hhsize]**

**replace pline = 120**

* Pour la reprise de la 1.4, les commandes suivantes seront nécessaire

**replace pgap = (pline-rpcinc)/pline if (rpcinc < pline)**

**sum pgap [aw=hhsize]**

**ifgt rpcinc, pline(100) alpha(1) hsize(hhsize)**

**Exercice 2 (3%)**

* 1. À l'aide du fichier data\_1, estimez les dépenses moyennes par équivalent adulte sans utiliser le poids de sondage et en utilisant la commande DASP **imean**. À quoi réfère cette statistique?

**R :**

* 1. Supposez différents cas d'initialisation du plan d'échantillonnage
* CAS1: Seulement en utilisant la variable *strata* pour initialiser la variable de stratification de la population échantillonnée.
* CAS2 : Seulement en utilisant la variable *psu* pour initialiser la variable d'unité primaire d’échantillonnage (primary sampling unit, PSU).
* CAS3: En utilisant la variable *strata* et *psu.*
* CAS4: En utilisant la variable *strata, psu* et la variable de poids de sondage*.*

Pour chacun de ces quatre cas, estimez les dépenses moyennes par équivalent adulte et donnez quelques explications sur le niveau des erreurs-types par rapport à celui de la question 1.1 et à ceux des autres cas.

**R :**

* 1. Vérifiez si les dépenses moyennes par équivalent adulte dans la région 1 sont supérieures au double de celles de la région 3. Discutez brièvement ce résultat.

**R :**

* 1. À l'aide de la commande DASP ***dimean***, évaluez si les dépenses moyennes par équivalent adulte pour les chefs de famille hommes sont plus élevées que celles des femmes chefs de famille. Discutez brièvement ce résultat.

**R :**

### Exercice 3 (5.5%)

**Q 3.1** Utilisez le fichier de données data\_1.dta, puis calculez la taille de la population des ménages échantillonnés.

**R :**

**Q 3.2** Ordonnez les dépenses par habitant en ordre croissant et générez ensuite la variable part de population (*ps*) qui comprend la proportion de la population échantillonnée avec les dépenses par habitant correspondantes. Sur cette base, générez les variables centiles (*p*) et quantiles (*q*).

**R :**

**Q 3.3** Dessinez la courbe de distribution cumulative (Axe X: les centiles et axe Y: les dépenses par habitant correspondantes) (domaine des centiles: min = 0 et max = 0,95).

**R :**

**Q 3.4** Tracez la courbe des quantiles (Axe X: centiles et axe Y: quantiles) (domaine des centiles: min = 0 et max = 0,95), et commentez brièvement les résultats.

**R :**

**Q 3.5** En utilisant DASP, dessinez la courbe des quantiles pour chacune des régions rurales et urbaines (domaine des centiles : min = 0 et max = 0,95), et discutez brièvement des résultats.

**R :**

**Q 3.6** À l'aide de DASP, dessinez les courbes de densité des dépenses par habitant en fonction du sexe du chef de ménage (domaine des dépenses par habitant: min = 0 et maximum = 1000000) et discuter brièvement des résultats.

**R :**