**Exercices semaines 1 et 2**

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 - Prénom, Nom" et veuillez les soumettre par la boîte de dépôt du portail de cours avant mardi le 2 février 23h59 (heure du Québec).

**Exercice 1 (4%)**

Supposons que la population est composée de 10 ménages vivant dans les régions A, B et C.

input region income hhsize

1 A 310 4

2 A 460 6

3 A 300 5

4 A 220 3

5 B 560 2

6 B 400 4

7 C 140 3

8 C 250 2

9 C 340 2

10 C 220 2

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

R : **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 : **sum pcinc [aw=hhsize]**

**sum (income)**

Q 1.3: En supposant que le seuil de pauvreté est égal à 120, générez la variable intensité de la pauvreté par habitant (pgap), puis estimez sa moyenne (l'intensité de la pauvreté par habitant doit être normalisée par le seuil de pauvreté).

R : **gen pline = 120**

**gen pgap = 0**

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

**sum pgap [aw=hhsize]**

Q 1.4: Refaites la question Q 1.3 en utilisant DASP.

R : **ifgt pcinc, pline(120) alpha(1) hsize(hhsize)**

Q 1.5: Supposons que le pouvoir d'achat dans la région B soit supérieur de 20% à celui de la région A et que celui de la région C soit supérieur de 40% 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 revenu réel par habitant (rpcinc).

R : **gen deflator = 1**

**replace deflator = 1.2 if region == 2**

**replace deflator = 1.4 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 110.

R : **sum rpcinc [aw=hhsize]**

**replace pline = 110**

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

**sum pgap [aw=hhsize]**

**ifgt rpcinc, pline(110) alpha(1) hsize(hhsize)**

**Exercice 2 (3%)**

2.1 À l'aide du fichier data\_2, estimez les dépenses moyennes par équivalent adulte sans utiliser le poids de sondage et en utilisant la commande DASP imean. À quoi cette statistique réfère-t-elle ?

R : **imean ae\_exp**

**L’équivalent adulte correspond à un ajustement facilitant la comparaison des dépenses des ménages vivant dans des contextes assez différents.**

2.2 En utilisant les variables strata, psu et la variable de poids de sondage, initialisez le plan d'échantillonnage, puis estimez la dépense moyenne par équivalent adulte.

R :

2.3 Vérifiez si la dépense moyenne par équivalent adulte dans la région 1 est supérieure au double de celle de la région 3.

R :

2.4 En utilisant la commande DASP dimean, vérifiez si la dépense moyenne par équivalent adulte pour les chefs de ménage hommes est plus élevée que celle des ménages dirigés par des femmes. Discutez brièvement vos résultats.

**Exercice 3 (5.5%)**

Q 3.1 Utilisez le fichier de données data\_2.dta, puis calculez la taille de la population des ménages échantillonnés.

R : **sum hsize**

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 avec les dépenses par habitant correspondantes. Sur cette base, générer les variables centiles (p) et quantiles (q).

R : - pour ordonner les dépenses par habitant en ordre croissant : **sort ae\_exp**

* Pour générer la part de la population (ps) :

**sum hsize**

**gen ps = hsize/r(sum)**

* Pour générer les variables centiles et quantiles :

**gen p = sum(ps)**

**gen q = ae\_exp**

**list, sep(0)**

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

R : **line p ae\_exp, title (The cumulative distribution curve) xtitle(dépenses ménages (y)) ytitle(F(y));**

Q 3.4 Tracez la courbe quantile (centiles sur l'axe X (0 à 0,95) et quantiles sur l'axe Y), et discutez brièvement les résultats.

R : **line q p , title(The quantile curve) xtitle(the percentile (p)) ytitle(The quantile Q(p));**

Q 3.5 En utilisant DASP, dessinez les courbes quantiles selon le sexe de la tête du ménage (centiles (0 à 0,95)), et discutez brièvement les résultats.

R : **c\_quantile ae\_exp, hsize(hhsize);**

Q 3.6 À l'aide du DASP, dessinez les courbes de densité des dépenses par habitant pour chacune des régions rurales et urbaines (domaine des dépenses par habitant : min = 0 et maximum = 1000000), et discutez brièvement des résultats.

R :

**histogram ae\_exp [fweight = hhsize], width(6)**

**start(0) kdensity kdenopts(width(3) gaussian);**

**cdensity ae\_exp , hs(hhsize) band(3) min(0)**

**max(1000000);**