Master 2

Théorie des sondages

Exercice 3

1. Les données suivantes proviennent d'un échantillon epsem à deux degrés de a = 10 grappes d'une population de N = 3 048 adultes dans A = 78 grappes de taille inégale, où y_α, est le total de la grappe pour l'indice de masse corporelle (IMC), et z_α est le nombre de personnes à qui un médecin a déjà dit qu'elles avaient un taux de cholestérol élevé parmi les adultes x_α sélectionnés dans chaque groupe. La conception est epsem car le taux de sondage au premier degré a été fixe $\frac{a}{A} = \frac{10}{68}$ et le taux au deuxième degré est de $\frac{b}{B} = \frac{10}{45}$. Par conséquent, la taille de l'échantillon est une variable aléatoire avec une taille d'échantillon attendue de 99,61.

α	y_α	z_α	x_α
1	340.25	2	12
2	192.5	3	7
3	100.15	3	4
4	281.43	2	10
5	316.85	3	12
6	344.36	1	13
7	281.78	3	10
8	304.86	5	10
9	243.26	3	8
10	337.86	3	11
Total	2,743.30	28	97

Sur la base de cet échantillon, répondez aux questions suivantes :

a) Calculer l'indice de masse corporelle (IMC) moyen, son erreur standard et son intervalle de confiance à 95 %
b) Estimez la proportion d'adultes à qui un médecin a déjà dit qu'ils avaient un taux de cholestérol élevé et son erreur type.
c) L'approximation de la série de Taylor est-elle adéquate pour les erreurs standard calculées en (a) et (b) ?
d) Calculez l'effet de conception et roh pour la proportion en (b).

2. Le tableau suivant est un résumé des résultats d'un échantillon aléatoire stratifié à allocation égale de taille n = 48 parmi les N = 3 048 adultes (N_h est la taille de la population de la strate, $\bar{y}_h$ est la moyenne de l'échantillon de la strate de l'indice de masse corporelle (IMC), s_h^2 est la variance des éléments de l'échantillon de la strate pour l'IMC et $\bar{z}_h$ est la proportion de l'échantillon de la strate d'adultes à qui un médecin a déjà dit qu'ils avaient un taux de cholestérol élevé) :

Stratum	N_h	$\bar{y}_h$	s_h^2	$\bar{z}_h$
1	467	25.53	28.663	0.125
2	388	25.07	34.779	0.250
3	479	28.23	21.477	0.375
4	446	28.65	23.348	0.250
5	570	30.04	67.386	0.250
6	698	33.76	36.500	0.375

a) Estimez l'IMC moyen ($\bar{y}$), son erreur standard et son intervalle de confiance à 95 %.

b) Estimez la proportion d'adultes à qui un médecin a déjà dit qu'ils avaient un taux de cholestérol élevé ($\bar{z}$) et son erreur type.

c) Calculez les effets de grappe pour la variance de l'IMC moyen (CONSEIL : vous aurez besoin d'une estimation SAS de s^2) et la proportion d'adultes à qui un médecin a déjà dit qu'ils avaient un taux de cholestérol élevé. Expliquez en une phrase ou deux pourquoi ces effets de grappe estimés sont différents.

d) Quelle est la répartition proportionnelle de n = 48 entre les strates, l'erreur type de l'IMC moyen sous cette répartition et l'effet de conception de la variance de la moyenne de l'échantillon réparti proportionnellement ? Qu'en est-il de la proportion d'adultes à qui un médecin a déjà dit qu'ils avaient un taux de cholestérol élevé ? L'allocation est-elle différente pour chaque variable ? Expliquer.

e) Sur la base des informations du tableau, quelle est l'allocation de Neyman de n = 48 dans les strates, l'erreur type de l'IMC moyen sous cette allocation et l'effet de conception de la variance de la moyenne de l'allocation de Neyman ?

f) En utilisant la notation de Kish pour l'élément de coût dans chaque strate, supposons qu'ils étaient J1 = J2 = J3 = J4 = 100 FCFA et J5 = J6 = 1000 FCFA. Supposons aussi que C – C0 = 30000. Calculez la répartition qui minimise la variance de l'IMC moyen pour ce coût total. Ne dépassez pas le budget de 30000. Estimez l'erreur type et l'effet de grappe de la variance de la moyenne sous cette allocation.