Serie 4 : Sondage à 2 degrés

Exercice 1 :

$\sum_{h\in E} A_h = 185$ $\quad \sum_{h\in E} A_h^2 = 1263$ $\quad N = 39.800$ clients

$n = 3980$ Agences

$m = 40$ Agences

$N_h = 10$ Clients $\forall h$ (taille d'une grappe)

1/ $\pi_{1h} = \frac{m}{n} = \frac{40}{3980} = \frac{2}{199}$ La probabilité de choisir une Agence

Puis on dénombre chaque agence $\Rightarrow$ Sondage par grappe où la grappe désigne une agence

2/ $P = \frac{1}{N} \cdot \sum_{h=1}^{3980} \sum_{i\in S_h} 1\!\!1_{\{c = oui\}}$ La proportion des clients ayant octroyé un crédit auprès de la banque

A_h : nombre des clients qui ont des prêts

$A_h = \sum_{i\in h} 1\!\!1_{\{c=oui\}}$ $\quad P = \frac{1}{N} \cdot \sum_{h=1}^{3980} A_h \Rightarrow \hat{P}_{grap} = \frac{1}{N} \cdot \sum_{h\in E_1} \frac{A_h}{\pi_{1h}}$

$\hat{P}_{grap\ por} = 0{,}46$

3/
$$Var(\hat{P}_{grap\ por}) = Var\left(\frac{1}{N} \cdot \sum_{h\in E_1} \frac{A_h}{\pi_{1h}}\right)$$
$$= \frac{1}{N^2} \cdot Var\left(\sum_{h\in E_1} \frac{A_h}{\pi_{1h}}\right)$$
$$= \frac{1}{N^2} \cdot \frac{n}{m} \cdot \frac{n-m}{n-1} \sum_{h=1}^{n} \left(A_h - \frac{\sum_{h=1}^{n} A_h}{n}\right)^2$$
$$= \frac{1}{(39800)^2} \cdot \frac{3980}{40} \cdot \frac{3980-40}{3980-1} \sum_{h=1}^{3980} \left(A_h - \frac{\sum_{h=1}^{3980} A_h}{3980}\right)$$
$$= 24 \cdot 10^{-5} \left[\sum_{h=1}^{3980} \left(A_h - \frac{\sum_{h=1}^{3980} A_h}{n}\right)^2\right] \cdot \frac{1}{n-1}$$
$$\Rightarrow \hat{V}ar(\hat{P}_{grap\ por}) = 24 \cdot 10^{-5} \cdot \frac{1}{m-1} \cdot \left[\sum_{h=1}^{40} \left(A_h - \frac{\sum_{h=1}^{40} A_h}{m}\right)^2\right]$$
$$= \frac{24 \cdot 10^{-5}}{39} \sum_{h=1}^{40} \left(A_h^2 + \left(\frac{\sum A_h}{m}\right)^2 - 2A_h \cdot \frac{\sum A_h}{m}\right)$$
$$= \frac{24 \cdot 10^{-5}}{39} \cdot \left(\sum A_h^2 + \sum \left(\frac{\sum A_h}{m}\right)^2 - 2 \cdot \sum \left(A_h \cdot \frac{\sum A_h}{n}\right)\right)$$
$$= \frac{24 \cdot 10^{-5}}{39} \left(1263 + 40 \times \left(\frac{185}{3980}\right)^2 - 2 \cdot \frac{(185)^2}{3980}\right) = 7{,}87 \times 10^{-3}$$

0,086425

$$IC_{95\%}(p) = \left[\hat{p}_{grap\,pair} \pm 1{,}96\,\sigma_{\hat{p}_{grap\,pair}}\right] = [0{,}46 \pm 0{,}17]$$

$$= [0{,}29\,;\,0{,}63]$$

Exercice 2 :

1/ Il s'agit d'un sondage aléatoire à 2 degrés, les unités primaires sont les 5 collèges tirés parmi les 50 selon un SAS sans remise à probabilités égales $\pi_{Ic} = \frac{5}{50} = \frac{1}{10}$. Les unités secondaires sont, pour chaque collège, les 10 collégiens de 7ème tirés selon un SAS sans remise à probabilité égale $\pi_{i|c}$ pour le collège c.

2/ Soit π_i la probabilité que l'individu i soit dans l'échantillon

$\pi_i = \pi_{i|c} \cdot \pi_{Ic}$

$\pi_{i|1} = \frac{10}{40} = \frac{1}{4}$ $\quad \pi_{i|2} = \frac{10}{20} = \frac{1}{2}$ $\quad \pi_{i|3} = \frac{10}{60} = \frac{1}{6}$ $\quad \pi_{i|4} = \frac{10}{40} = \frac{1}{4}$

$\pi_{i|5} = \frac{10}{48} = \frac{5}{24}$

$$\hat{T} = \sum_{c \in E_I} \sum_{i \in E_c} \frac{y_i}{\pi_i} = \sum_{c \in E_I} \sum_{i \in E_c} \frac{y_i}{\pi_{i|c} \cdot \pi_{Ic}} = \sum_{c \in E_I} \frac{1}{\pi_{Ic}} \sum_{i \in E_c} \frac{y_i}{\pi_{i|c}}$$

$$\hat{T} = 10 \sum_{c \in E_I} \hat{t}_c$$

$$= 10(40 \times 12 + 20 \times 8 + 60 \times 10 + 40 \times 12 + 48 \times 11) = 22.480$$

3/ $\hat{N} = \frac{1}{\pi_{Ic}} \sum_{c \in E_I} N_c = \frac{50}{5}(40 + 20 + 60 + 40 + 48) = 2080$

4/ $N = 2000$

$$\hat{\bar{Y}} = \frac{1}{2000} \hat{T} = \frac{22480}{2000} = 11,24$$

$m = 50$

La moyenne sur l'échantillon est $\hat{\bar{Y}}_{ech} = \frac{1}{5}(12 + 8 + 10 + 12 + 11) = 10,6 < \hat{\bar{Y}}$

$\hat{\bar{Y}}_{ech}$ n'est généralement pas un bon estimateur de la moyenne de la population

5/ $Var(\hat{T}) = Var(U_P) + Var(U_S)$

$$Var(U_P) = \sum_{c=1}^{50} \frac{T_c^2}{\pi_{Ic}^2}(\pi_{Ic} - \pi_{Ic}^2) + \sum_{c=1}^{50} \sum_{\substack{k=1 \\ k \neq c}}^{50} \frac{T_c}{\pi_{Ic}} \cdot \frac{T_k}{\pi_{Ik}} (\pi_{Ick} - \pi_{Ic} \cdot \pi_{Ik})$$

$$\pi_{Ick} = \frac{5}{50} \cdot \frac{4}{49} = \frac{2}{245}$$

$$Var(U_P) = \sum_{c=1}^{50} 10 \cdot \frac{9}{10} T_c^2 + \sum_{c=1}^{50} \sum_{\substack{k=1 \\ k \neq c}}^{50} 10^2 \cdot T_c \cdot T_k \left(\frac{5}{50} \cdot \frac{4}{49} - \frac{1}{10^2}\right) \frac{1}{5} \cdot \frac{2}{49}$$

$$= 9\sum_{c=1}^{50} T_c^2 - \frac{45}{245} \cdot \sum_{c=1}^{50}\sum_{\substack{k=1\\k\neq c}}^{50} T_c.T_k$$

$$Var(U_s) = \sum_{c=1}^{50} 10 Var(\hat{T}_c) = 10. \sum_{c=1}^{50} (1-f_c) \frac{\sigma_{cc}^2}{m_c}$$

$m_c = 10$

$$Var(U_s) = 10. \sum_{c=1}^{50} \left(1 - \frac{10}{N_c}\right) \sigma_{cC}^2$$

Oma alors

$$Var(\hat{T}) = 9\sum_{c=1}^{50} T_c^2 - \frac{45}{245} \sum_{c=1}^{50}\sum_{\substack{k=1\\k\neq c}}^{50} T_c.T_k + 10\sum_{c=1}^{50}\left(1 - \frac{10}{N_c}\right)\sigma_{cC}^2$$

6/ $\hat{Var}(\hat{T}) = 12891408$

$$\hat{Var}(\hat{\bar{Y}}) = \frac{1}{N^2}\hat{Var}(\hat{T}) = \frac{12891408}{2000^2} = 3,22$$

$$IC_{1\alpha}(\bar{Y}) = \left[\hat{\bar{Y}} \pm 1.96\sqrt{3,22}\right]$$