Ecole Supérieure De La Statistique Et De L'Analyse De L'Information

Enquête Sur La Consommation En Eau Minérale En Tunisie

Réalisé par :
YAHYAOUI Marwa
BEDHIEFI Imen
NEFZAOUI Aida

Table des matières :

Introduction à la théorie des sondages :

La théorie des sondages s’est développée à partir des années 1930 dans le monde anglo-saxon ainsi qu’en Inde. Bien que reposant sur les mêmes principes que la statistique mathématique classique, elle en diffère sensiblement dans son esprit en raison d’objectifs spécifiques.

Cette théorie se consacre essentiellement au problème de la sélection de l’échantillon et à la recherche d’estimateurs.

Du fait qu’elle porte sur des populations finies, d’existence bien concrète, elle ne peut ignorer les contraintes du monde réel, ce qui n’est peut-être pas sans lien avec le faible intérêt qu’elle suscite chez nous.

Chapitre 1: les plans simples :

Ce chapitre traite exclusivement des plans simples qui forment souvent la base de plans de sondage plus complexes, tel que les plans stratifiés ou par grappes.

1.1 Plans simples sans remise :

1.1.1 Plans de sondage et probabilité d'inclusion :

Un plan de taille fixe n est dit simple sans remise si :

$$p(x) = \begin{cases} \binom{N}{n}^{-1}, & \text{si } |s| = n \\ 0, & \text{sinon} \end{cases}$$

avec $n \in 1, ...N$

Il est souvent utile pour nous statisticiens de connaître les probabilités d'inclusion afin de pouvoir établir le pi-estimateur et sa variance par exemple.

$$\pi_k = \sum_{s \ni k} p(s) = \underbrace{\binom{N-1}{n-1}}_{\substack{\text{nb. d'échantillons} \\ \text{contenant } k}} \binom{N}{n}^{-1} = \frac{(N-1)!}{(n-1)!(N-n)!} \frac{n!(N-n)!}{N!} = \frac{n}{N}$$

$$\pi_{k\ell} = \sum_{s:\, k,\ell \in s} p(s) = \underbrace{\binom{N-2}{n-2}}_{\substack{\text{nb. échantillons} \\ \text{contenant } k \text{ et } \ell}} \binom{N}{n}^{-1} = \frac{(N-2)!}{(n-2)!(N-n)!} \frac{n!(N-n)!}{N!} = \frac{n(n-1)}{N(N-1)}$$

Remarque. Notons que $\pi_{k\ell} \neq \pi_k \pi_\ell$, indiquant une dépendance entre les unités choisies dû au tirage sans remise.

Des deux expressions précédentes, on en déduit

$$\Delta_{k\ell} = \begin{cases} \frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2} = -\frac{n(N-n)}{N^2(N-1)}, & k \neq \ell, \\ \frac{n}{N}\left(1 - \frac{n}{N}\right) = \frac{n(N-n)}{N^2}, & k = \ell. \end{cases}$$

1.1.2 Le -estimateur pour ces plans :

Le π-estimateur pour une moyenne muy est :

$$\hat{\mu}_{y,\pi} = \frac{1}{N} \sum_{k \in \mathcal{U}} \frac{y_k}{\pi_k} 1_{\{k \in S\}} = \frac{N}{nN} \sum_{k \in \mathcal{U}} y_k 1_{\{k \in S\}} = \overline{y},$$

et le π–estimateur du total t_y est évidemment $\hat{t}_{y,\pi} = N\overline{y}$, avec

$$\overline{y} = \frac{1}{n} \sum_{k \in \mathcal{U}} y_k 1_{\{k \in S\}}.$$

Rappelons que puisque le plan est à taille fixe et que les probabilités d'inclusion des deux premiers ordres sont strictement positives, on peut utiliser la formule de la variance de muy,pi; trouvés par Sen-Yates-Grundy :

$$\begin{aligned} \operatorname{Var}(\hat{\mu}_{y,\pi}) &= -\frac{1}{2N^2} \sum_{\substack{k,\ell \in \mathcal{U} \\ k \neq \ell}} \left(\frac{y_k}{\pi_k} - \frac{y_\ell}{\pi_\ell}\right)^2 \Delta_{k\ell} = \frac{1}{2N^2} \times \frac{N-n}{n(N-1)} \sum_{\substack{k,\ell \in \mathcal{U} \\ k \neq \ell}} (y_k - y_\ell)^2 \\ &= \frac{N-n}{nN} S_y^2, \end{aligned}$$

avec

$$S_y^2 = \frac{1}{N-1} \sum_{k \in \mathcal{U}} (y_k - \mu_y)^2 = \frac{1}{2N(N-1)} \sum_{\substack{k,\ell \in \mathcal{U} \\ k \neq \ell}} (y_k - y_\ell)^2.$$

Pour un plan de taille fixe n, simple et sans remise la variance corrigée de la population S2y est estimée sans biais par :

$$\widehat{S_y^2} = \frac{1}{n-1}\sum_{k\in\mathcal{U}}(y_k - \bar{y})^2 1_{\{k\in S\}}.$$

On peut aussi estimer sans biais la variance de muy,pi par

$$\widehat{\mathrm{Var}}(\hat{\mu}_{y,\pi}) = \frac{N-n}{N}\frac{\widehat{S_y^2}}{n},$$

Et pour le pi-estimateur du total ty,pi par

$$\widehat{\mathrm{Var}}(\hat{t}_{y,\pi}) = N(N-n)\frac{\widehat{S_y^2}}{n}.$$

1.2 Plan simple avec remise :

Le plan de taille fixe n simple et avec remise correspond au cadre de variable aléatoire indépendante et identiquement distribuées de moyenne

$$\mu_y = \frac{1}{N}\sum_{k\in U} y_k$$

Et de variance

$$\sigma_y^2 = \frac{1}{N}\sum_{k\in\mathcal{U}}(y_k - \mu_y)^2$$

La moyenne sur la population y est estimée sans biais par

$$\hat{\mu}_y = \frac{1}{n}\sum_{k\in\mathcal{U}} y_k 1_{\{k\in S\}} = \bar{y}$$

De plus, puisque les individus de l'échantillon sont sélectionnés indépendamment et sont de même loi,

$$\mathrm{Var}(\hat{\mu}_y) = \frac{\sigma_y^2}{n}$$

La variance de y est estimée sans biais par

$$\widehat{\mathrm{Var}}(\hat{\mu}_y) = \frac{\widehat{\sigma_y^2}}{n}.$$

1.3 Comparaison des plans simples avec et sans remise :

Le sondage simple et sans remise est toujours préférable à celui avec remise. En effet si l'on appelle muypi^ et muypi~ les -estimateurs de la moyenne avec et sans remise, alors pour tout n >=2

$$\frac{\mathrm{Var}(\tilde{\mu}_{y,\pi})}{\mathrm{Var}(\hat{\mu}_{y,\pi})} = \frac{(N-n)}{N}\times\frac{S_y^2}{\sigma_y^2} = \frac{N-n}{N}\times\frac{N}{N-1} = \frac{N-n}{N-1} < 1$$

Chapitre 2 :
Plans a probabilités inégales :

Ce chapitre explique comment nous pouvons bénéficier de la connaissance d'un caractère auxiliaire pour obtenir des estimations plus précises. Le but du sondage est de maximiser la précision de l'estimateur, c'est à dire, minimiser la variance. Comme on a vu dans l'expression de celle du -estimateur dans le cas du sondage aléatoire simple, ceci est possible si on prend des probabilités d'inclusion proportionnelles aux individus Yi. Mais ceci impossible car les Yi sont inconnues. Pour cela on recourt au sondage à probabilités inégales. Ce plan consiste à trouver une variable auxiliaire correlée positivement à la variable d'intérêt et pour laquelle les probabilités d'inclusion y sont proportionnelles. Plus précisément :

$$\pi_k = \frac{x_k n}{\sum_{l \in P} x_l} \text{ pour k} \in \text{ P et l} \neq \text{k.}$$

L'estimateur de la moyenne est :

$$\overline{y} = \frac{1}{nN} \sum_{i \in E} \frac{y_i}{p_i}$$

La variance de l'estimateur de la moyenne est :

$$var\left(\overline{y}\right) = \frac{1}{N^2} Var\left(\frac{1}{n} \sum_{i \in E} \frac{y_i}{p_i}\right) = \frac{1}{n^2 N^2} \sum_{i \in E} Var\left(\frac{y_i}{p_i}\right) = \frac{1}{nN^2} Var\left(\frac{y_1}{p_1}\right)$$

2.1 Sondage à probabilité inégale avec remise :

On définit :

$$p_i = \frac{x_i}{\sum_{i \in p} x_i} \text{ et : } v_i = \sum_{j=1}^{i} p_j$$

On génère une variable aléatoire continue uniforme u sur [0,1] et on sélectionne l'unité i telle que

$$v_{i-1} \leq u \leq v_i$$

On répète n fois cette opération de manière indépendante pour obtenir l'échantillon de taille n.

2.2 Sondage à probabilité inégale sans remise :

Chaque tirage modifie les conditions des tirages suivants, il existe plusieurs méthodes de tirages sans remise proportionnellement à une variable auxiliaire. Ces probabilités d'inclusion sont comprises entre 0 et 1. Comme la taille de l'échantillon est fixe n,ces probabilités respectent l'égalité suivante :

$$\sum_{i}^{N} \pi_i = n$$

L'estimateur de Horvitz-Thompson de la moyenne pour un sondage a probabilités inégales sans remise est défini par :

$$\widehat{Y}_{PISR} = \sum_{i=1}^{N} \frac{X_i}{\pi_i}$$

Chapitre 3 : Stratification :

La technique de stratification est largement utilisée en sondage car elle permet facilement d'introduire de l'information auxiliaire pour la construction d'un plan de sondage adequat.

3.1 Population et strates :

Supposons que la population U soit partitionnée en H sous-ensembles U1,...,UH appelés strates et tels que

$$\bigcup_{i=1}^{H} \mathcal{U}_i = \mathcal{U}, \qquad \mathcal{U}_i \cap \mathcal{U}_h = \emptyset, \quad i \neq h$$

Chaque strate UH admet une taille Nh et l'on a bien évidemment

$$\sum_{h=1}^{H} N_h = N$$

où N est la taille de la population U.
Remarque : Les tailles des strates Nh sont ici supposées connues et constituent l'information

Le total sur la population s'écrit :

$$t_y = \sum_{k \in \mathcal{U}} y_k = \sum_{h=1}^{H} \sum_{k \in \mathcal{U}_h} y_k = \sum_{h=1}^{H} t_{y,h}$$

La moyenne sur la population s'écrit :

$$\mu_y = \frac{1}{N}\sum_{k\in\mathcal{U}} y_k = \frac{1}{N}\sum_{h=1}^{H}\sum_{k\in\mathcal{U}_h} y_k = \frac{1}{N}\sum_{h=1}^{H} N_h\mu_{y,h}.$$

On définit également la variance et la variance corrigée sur une strate UH par

$$\sigma^2_{y,h} = \frac{1}{N_h}\sum_{k\in\mathcal{U}_h}(y_k-\mu_{y,h})^2.$$

et

$$S^2_{y,h} = \frac{1}{N_h-1}\sum_{k\in\mathcal{U}_h}(y_k-\mu_{y,h})^2$$

Remarque : La variance sur la population sigma2y s'écrit

$$= \sigma^2_{y,\text{intra}} + \sigma^2_{y,\text{inter}}$$

avec

$$\sigma^2_{y,\text{intra}} = \frac{1}{N}\sum_{h=1}^{H} N_h\sigma^2_{y,h}$$

et

$$\sigma^2_{y,\text{inter}} = \frac{1}{N}\sum_{h=1}^{H} N_h(\mu_{y,h}-\mu_y)^2$$

3.2 Echantillons, probabilités d'inclusion et estimation :

Un sondage est dit stratifié si, pour chaque strate, on tire un échantillon selon un sondage aléatoire simple sans remise de taille xe nh et que les tirages au sein de chaque strate sont mutuellement indépendant. Soit Sh l'échantillon aléatoire tiré dans la strate Uh a l'aide d'un plan de sondage ph .
L'échantillon aléatoire S obtenu au nal est donc

$$S = \bigcup_{h=1}^{H} S_h$$

Le plan de sondage associé n'est rien d'autre que

$$p(s) = \prod_{h=1}^{H} p_h(s_h), \qquad s = \bigcup_{h=1}^{H} s_h$$

et la taille de l'échantillon S est

$$n = \sum_{h=1}^{H} n_h$$

Pour les probabilités d'inclusion d'ordre un et si l'unité k appartient a la strate Uh, alors

$\pi_k = \frac{n_h}{N_h}$ puisqu'on a effectué un plan simple sans remise de taille nh pour cette strate.
Pour les probabilités d'inclusion d'ordre deux : Si k et l appartiennent a la même strate Uh alors :
Si k et l appartiennent a deux strates différentes Uh1 et Uh2 alors (par indépendance entre les strates)

$$\pi_{k\ell} = \frac{n_h(n_h-1)}{N_h(N_h-1)}$$

$$\pi_{kl} = \pi_k \pi_\ell = \frac{n_{h_1}}{N_{h_1}} \frac{n_{h_2}}{N_{h_2}}$$

$$\Delta_{k\ell} = \begin{cases} \frac{n_h}{N_h}\left(1 - \frac{n_h}{N_h}\right), & k = \ell,\ k \in \mathcal{U}_h, \\ -\frac{n_h(N_h-n_h)}{N_h^2(N_h-1)}, & k \neq \ell,\ k, \ell \in \mathcal{U}_h, \\ 0, & k \in \mathcal{U}_h,\ \ell \notin \mathcal{U}_h. \end{cases}$$

Par conséquent on a :

Du coup les ?-estimateurs du total ty et de la moyenne muy sont

$$\hat{t}_{y,\text{strat}} = \sum_{k\in S} \frac{y_k}{\pi_k} = \sum_{h=1}^{H} \sum_{k\in S_h} \frac{N_h y_k}{n_h} = \sum_{h=1}^{H} \hat{t}_{y,h},$$

et

$$\hat{\mu}_{y,\text{strat}} = \frac{1}{N} \sum_{h=1}^{H} \frac{N_h}{n_h} \sum_{k\in S_h} y_k = \frac{1}{N} \sum_{h=1}^{H} N_h \bar{y}_h,$$

Puisque les strates sont indépendantes, la variance de ces estimateurs se calcule facilement

$$\text{Var}(\hat{t}_{y,\text{strat}}) \overset{\text{ind}}{=} \sum_{h=1}^{H} \text{Var}(\hat{t}_{y,h}) \overset{\text{simple}}{=} \sum_{h=1}^{H} N_h(N_h - n_h)\frac{S_{y,h}^2}{n_h}$$

variance qui s'estime sans biais par

$$\widehat{\text{Var}}(\hat{t}_{y,\text{strat}}) = \sum_{h=1}^{H} N_h(N_h - n_h)\frac{s_{y,h}^2}{n_h},$$

avec

$$s_{y,h}^2 = \frac{1}{n_h - 1} \sum_{k\in S_h} (y_k - \bar{y}_h)^2, \qquad h = 1, \ldots, H.$$

3.3 Plan stratifié et allocation proportionnelle :

Un plan stratifié est dit à allocation proportionnelle si

$$\frac{n_h}{N_h} = \frac{n}{N}, \qquad h = 1, \ldots, H$$

Les -estimateurs du total et de la moyenne sont alors

$$\hat{t}_{y,\text{strat. prop.}} = \sum_{h=1}^{H} \hat{t}_{y,h} = \frac{N}{n} \sum_{k \in S} y_k$$

$$\hat{\mu}_{y,\text{strat. prop.}} = \frac{1}{n} \sum_{k \in S} y_k.$$

La variance du total est alors

$$= \frac{N-n}{n} \sum_{h=1}^{H} N_h S_{y,h}^2$$

Remarque : Lorsque les tailles des strates Nh sont suffisamment grandes, alors

$$S_{y,h}^2 \approx \sigma_{y,h}^2$$

Et donc

$$\text{Var}(\hat{t}_{y,\text{strat. prop.}}) \approx \frac{N-n}{n} \sum_{h=1}^{H} N_h \sigma_{y,h}^2 = N(N-n) \frac{\sigma_{y,\text{intra}}^2}{n}$$

Tandis que la variance du total pour un plan simple sans remise verie

$$\text{Var}(\hat{t}_{y,\pi}) \approx N(N-n) \frac{\sigma_y^2}{n}$$

On estime sans biais cette variance par

$$\widehat{\text{Var}}(\hat{t}_{y,\text{strat. prop.}}) = \frac{N-n}{n} \sum_{h=1}^{H} N_h \widehat{S_{y,h}^2},$$

avec

$$\widehat{S_{y,h}^2} = \frac{1}{n_h - 1} \sum_{k \in S_h} (y_k - \bar{y}_h)^2, \qquad h = 1, \ldots, H.$$

3.4 Plan stratifié optimal pour le total :

On cherche les tailles d'échantillon n1,...,nh minimisant la variance du pi-estimateur du total ty pour une taille d'échantillon fixe n, i.e., minimiser

$$\text{Var}(\hat{t}_{y,\text{strat}}) = \sum_{h=1}^{H} N_h (N_h - n_h) \frac{S_{y,h}^2}{n_h}$$

Sous la contrainte

$$\sum_{h=1}^{H} n_h = n.$$

et il vient

$$n_h = n \frac{N_h S_{y,h}}{\sum_{j=1}^{H} N_j S_{y,j}}, \qquad h = 1, \ldots, H$$

La formule précédente est assez instructive puisqu'elle indique qu'il faut surreprésenter les strates qui ont une forte variabilité. De plus il peut arriver également que nh > Nh pour un h dans 1,...,H. Pour de tels cas, on posera alors nh = Nh et on déterminera les tailles optimales sur les strates restantes. Supposons que nos tailles optimales soient des entiers et telles que nh < Nh pour tout h. Alors la variance du pi-estimateur est alors

$$= \frac{1}{n} \left(\sum_{h=1}^{H} N_h S_{y,h} \right)^2 - \sum_{h=1}^{H} N_h S_{y,h}^2$$

3.5 Prise en compte du coût :

On visera une allocation optimale pour un budget fixé C sous la contrainte

$$\sum_{h=1}^{H} n_h C_h = C$$

représente le coût d'interroger une unité dans la strate Uh .

Chapitre 4 : Plans par grappe et à plusieurs degrés :

Dans ce chapitre nous allons voir comment une variable auxiliaire peut être utilisée non pas pour ameliorer la precision de nos estimations mais le déroulement d'une enquête.

4.1 Plan par grappe :

Supposons que la population U soit partitionnée en M sous-ensembles U1,...,UM appelés grappes et tels que

$$\bigcup_{i=1}^{M} \mathcal{U}_i = \mathcal{U}, \qquad \mathcal{U}_i \cap \mathcal{U}_j = \emptyset, \quad i \neq j$$

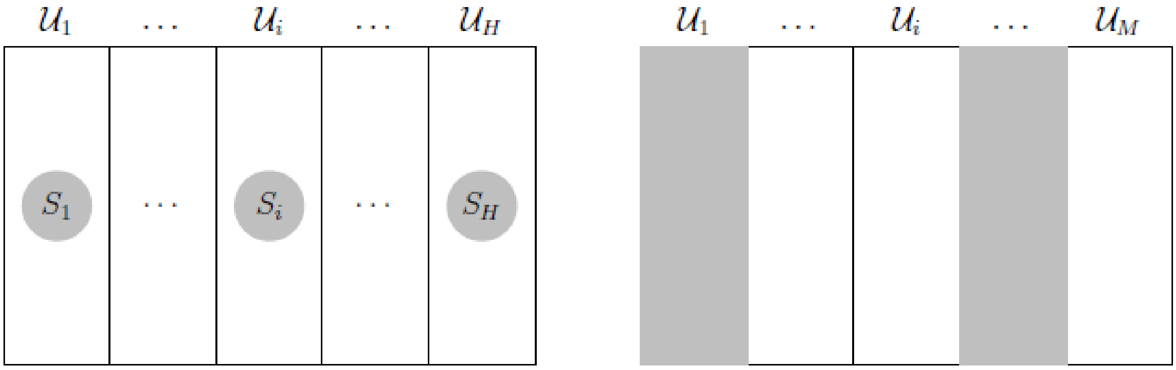
Le total de la population s'écrit :

$$t_y = \sum_{k \in \mathcal{U}} y_k = \sum_{i=1}^{M} \sum_{k \in \mathcal{U}_i} y_k = \sum_{i=1}^{M} t_{y,i}$$

De même la moyenne sur la population s'écrit :

$$\mu_y = \frac{1}{N} \sum_{k \in \mathcal{U}} y_k = \frac{1}{N} \sum_{i=1}^{M} \sum_{k \in \mathcal{U}_i} y_k = \frac{1}{N} \sum_{i=1}^{M} N_i \mu_{y,i}$$

C'est une illustration de la différence entre un plan de sondage stratifié (gauche) et par grappe (droite) :



Pour l'un, un échantillon aléatoire est prélevé dans chaque strate.Pour l'autre, un échantillon aléatoire sur les grappes est prélevé et chaque grappe ainsi piochée est entièrement retenue.
Un plan est dit par grappes si l'on procède comme suit :
1. On sélectionne un échantillon aléatoire de grappes Sg selon un plan de sondage pg définit sur les parties non vides de Ug = 1,...,M 2. Toutes les unités des grappes sélectionnées sont alors retenues.

Un échantillon aléatoire S issu d'un plan par grappes s'écrit

$$S = \bigcup_{i \in S_g} \mathcal{U}_i$$

et sa taille est

$$n_S = \sum_{i \in S_g} N_i$$

Probabilités d'inclusion d'ordre 1 :

$$\pi_k = \sum_{\substack{s \in \mathscr{S}_g \\ i \in s}} p_g(s) \stackrel{\text{def}}{=} \pi_{g,i}, \qquad k \in \mathcal{U}_i, \quad i \in \mathcal{U}_g$$

Avec est l'ensemble des échantillons possibles de probabilités d'inclusion d'ordre 2 :

$$\pi_{k\ell} = \begin{cases} \pi_{g,i}, & k, \ell \in \mathcal{U}_i \\ \pi_{g,ij}, & k \in \mathcal{U}_i,\ \ell \in \mathcal{U}_j, \end{cases}$$

avec

$$\pi_{g,ij} = \sum_{\substack{s \in \mathcal{S}_g \\ i,j \in s}} p_g(s), \qquad i, j \in \mathcal{U}_g, \quad i \neq j$$

La variance du π-estimateur du total ty est =

$$\sum_{i=1}^{M} \frac{t_{y,i}^2}{\pi_{g,i}^2} \pi_{g,i}(1 - \pi_{g,i}) + \sum_{i \neq j} \frac{t_{y,i} t_{y,j}}{\pi_{g,i} \pi_{g,j}} (\pi_{g,ij} - \pi_{g,i} \pi_{g,j})$$

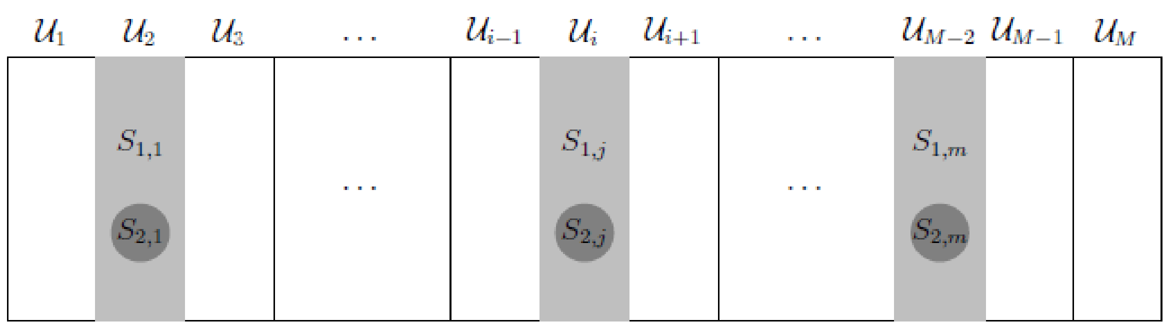
Si le plan est de taille fixe, on aura

$$\mathrm{Var}(\hat{t}_{y,\pi}) = -\frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{M} \left(\frac{t_{y,i}}{\pi_{g,i}} - \frac{t_{y,j}}{\pi_{g,j}} \right)^2 \Delta_{g,ij}$$

4.2 Plans à deux degrés :

Les plans a deux degrés portent bien leurs noms puisqu'ils consistent en un double échantillonnage :

1. sur les unités primaires
2. puis les unités secondaires



C'est une illustration du concept de plan a deux degrés. L'échantillon du premier degré est de taille m et est

$$S_1 = \cup_{j=1}^{m} S_{1,j}$$

L'échantillon “final” obtenu par un plan a deux degres est alors

$S = \cup_{j \in S_1} S_{2,j}$

Pour effectuer un plan à deux degrés, il faut donc construire un échantillon S1 d'unités primaires à partir d'un plan de sondage p1(.) sur 1, . . . ,M;
pour chaque unité primaire sélectionnée, construire un échantillon S2 sur les unité secondaires à partir d'un plan de sondage p2(.). Il est souhaitable que les plans à deux degrés possèdent les deux propriétés suivantes : Invariance : le plan du second degré p2(.) est indépendant du premier plan p1(.), i.e.,
Pr(S2 = s2 j S1 = s1) =Pr(S2 = s2) .

Indépendance : les tirages du second degré sont mutuellement indépendants. Probabilités d'inclusion d'ordre 1 et 2 :

$$\pi_{1,i} = \Pr(\mathcal{U}_i \in S_1)$$

$$\pi_{1,ij} = \Pr(\mathcal{U}_i \in S_1, \mathcal{U}_j \in S_1).$$

$\pi_{k|i}$ = la probabilité de sélectionner l'unité (secondaire) k sachant que l'unité (primaire) a été choisie

$\pi_{k\ell|i}$ = la probabilité d'inclusion d'ordre 2 sachant que a été retenue

$$\pi_k = \Pr(k \in S_{2,i}, i \in S_1) = \Pr(k \in S_{2,i} \mid i \in S_1)\Pr(i \in S_1) = \pi_{k|i}\pi_{1,i}$$

$$\pi_{k\ell} = \begin{cases} \pi_{k\ell|i}\pi_{1,i}, & k, \ell \in \mathcal{U}_i, \\ \pi_{k|i}\pi_{\ell|j}\pi_{1,ij}, & k \in \mathcal{U}_i,\ \ell \in \mathcal{U}_j,\ i \neq j \end{cases}$$

Le total :

$$t_y = \sum_{k \in \mathcal{U}} y_k = \sum_{i=1}^{M} \sum_{k \in \mathcal{U}_i} y_k = \sum_{i=1}^{M} t_{y,i}$$

Le ?-estimateur du total :

$$\hat{t}_{y,\pi} = \sum_{i \in S_1} \sum_{k \in S_{2,i}} \frac{y_k}{\pi_{k|i}\pi_{1,i}} = \sum_{i \in S_1} \frac{\hat{t}_{y,i}}{\pi_{1,i}}$$

Collecte des données :

Période d'enquête :
L'enquête s'est déroulée du 30/10/2016 au 15/11/2016

Enquêtrices :
Nefzaoui Aida
Yahyaoui Marwa
Bedhiefi Imen

Techniques d'enquête :
L'enquête s'est déroulée via l'application de création de questionnaire de Google docs et par partage sur le réseau social Facebook.

Traitement des données :
Les données ont été encodées avec EXCEL, converties au format .CSV et traitées avec le logiciel R.

La construction du questionnaire :
Le questionnaire a été rédigé avec la supervision de Mme Hela Ouaili-Mallek

Bases de données :
Les données collectées sont encodées dans une data. Frame de 101 ligne (100 individus interrogés) et de 7 colonnes chacune relative à une question.

Population cible :
Tout le monde

Population source :
L'entourage

Variable d'intérêt :
Moyenne d'eau consommée

Variables auxiliaires :
Sexe

- homme,
- femme

Age

- [18,30[
- [30,60[
- [60 et plus[

Sport :

- Non
- Une fois par semaine
- Plus de 2 fois

Maladies :

- Oui
- Non

Partie pratique :

L'enquête proposée a été conçue auprès de 100 tunisiens.
La consommation en eau minérale constitue la variable d'intérêt pour le reste de l'étude.
Ci-dessous le questionnaire associé :

Sexe : *

- Femme
- Homme

Age : *

- [18,30[
- [30,60[
- 60 et plus

Combien de litres d'eau buvez-vous par jour en moyenne ? *

- Moins d'un litre
- 1 à 2 litres
- Plus de 2 litres

Vous buvez uniquement de l'eau minérale ? *

- Oui
- Non

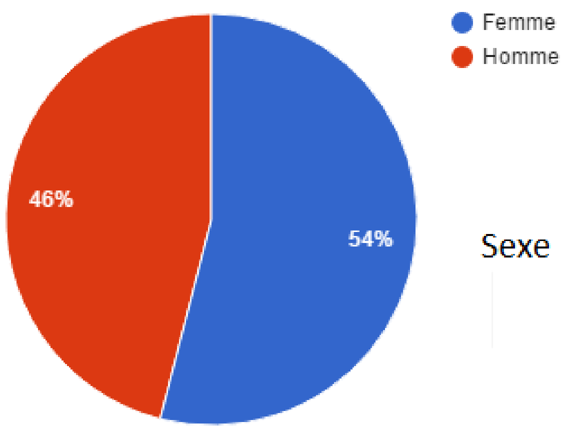
Souffrez-vous d'une maladie vous obligeant à boire de l'eau minérale ? *

- Oui
- Non

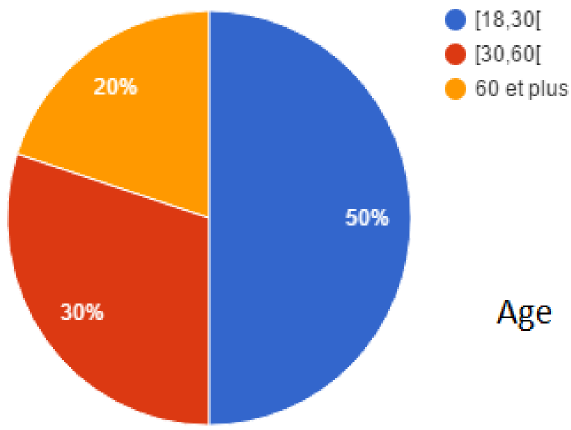
Vous pratiquez une activité physique ? *

- Non
- 1 fois par semaine
- Plus de 2 fois par semaine

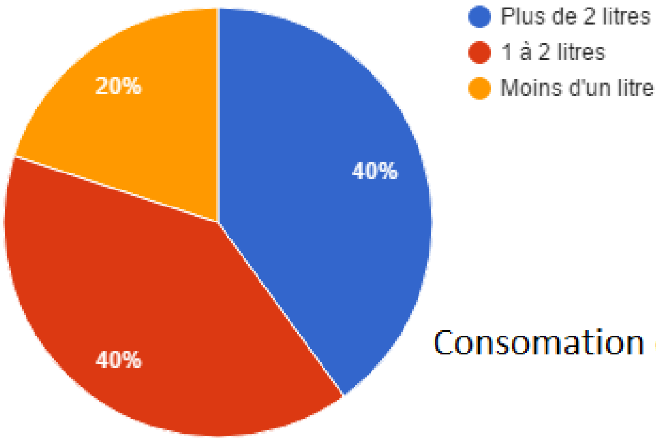
Statistique descriptive :



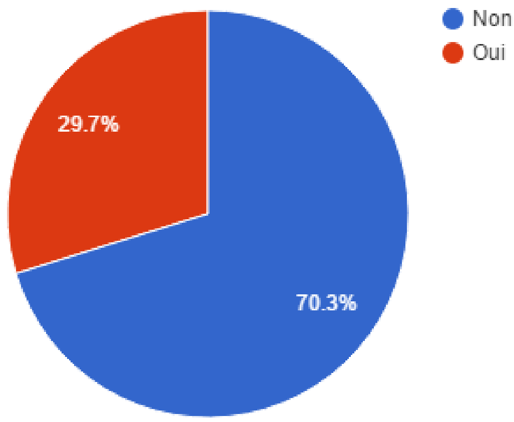
Sexe



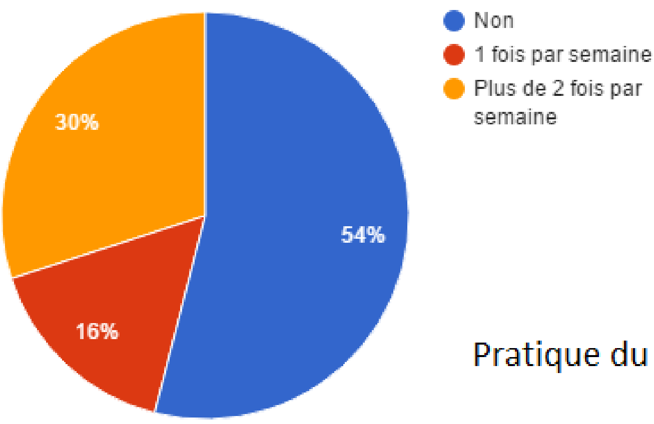
Age



Consomation en eau



Les personnes atteintes d'une maladie

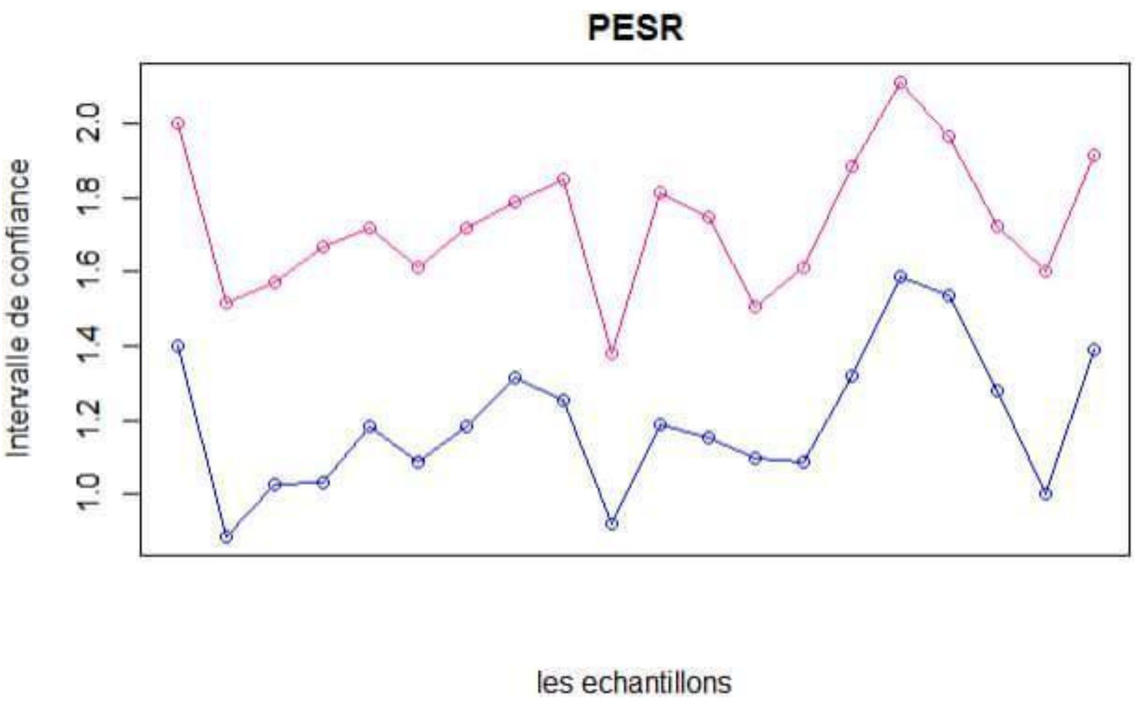


Pratique du sport

Principaux résultats:

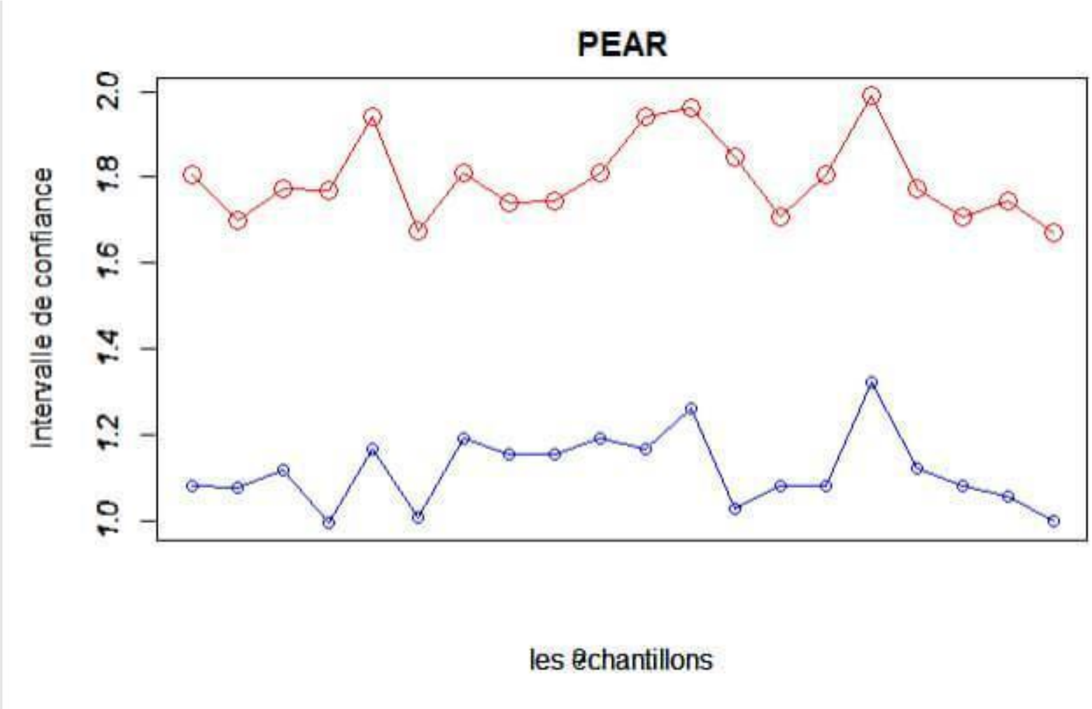
1. Sondage par PESR :

	ybarPESR	varPESR	icPESRmin	icPESRmax	longueurIC	moy_PESR
1	1.40	0.02484211	1.091077	1.708923	0.6178460	1.5625
2	1.40	0.02063158	1.118472	1.681528	0.5630569	1.5625
3	1.65	0.01800000	1.387038	1.912962	0.5259232	1.5625
4	1.60	0.01642105	1.348836	1.851164	0.5023270	1.5625
5	1.75	0.02473684	1.441732	2.058268	0.6165357	1.5625
6	1.65	0.01800000	1.387038	1.912962	0.5259232	1.5625
7	1.75	0.01210526	1.534353	1.965647	0.4312938	1.5625
8	1.75	0.01631579	1.499643	2.000357	0.5007144	1.5625
9	1.35	0.02221053	1.057897	1.642103	0.5842053	1.5625
10	1.65	0.02221053	1.357897	1.942103	0.5842053	1.5625
11	1.55	0.01884211	1.280958	1.819042	0.5380849	1.5625
12	1.80	0.02147368	1.512784	2.087216	0.5744330	1.5625
13	1.50	0.02526316	1.188470	1.811530	0.6230600	1.5625
14	1.60	0.02484211	1.291077	1.908923	0.6178460	1.5625
15	1.50	0.02105263	1.215613	1.784387	0.5687734	1.5625
16	1.45	0.01884211	1.180958	1.719042	0.5380849	1.5625
17	1.55	0.02305263	1.252411	1.847589	0.5951772	1.5625
18	1.45	0.01884211	1.180958	1.719042	0.5380849	1.5625
19	1.35	0.01378947	1.119840	1.580160	0.4603201	1.5625
20	1.55	0.01884211	1.280958	1.819042	0.5380849	1.5625



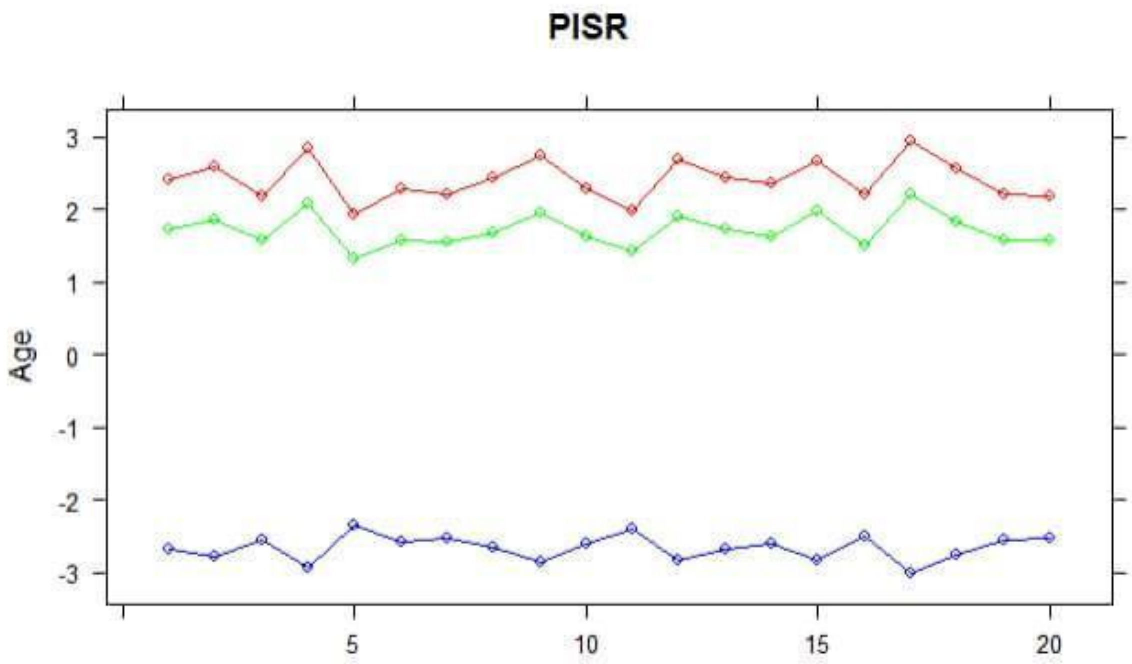
2. Sondage par PEAR :

	ybarPEAR	varPEAR	icPEARmin	icPEARmax	longueurIC	moy_PEAR
1	1.342105	-0.02901298	1.0082546	1.675956	0.6677013	1.466821
2	1.552632	-0.02230646	1.2598988	1.845364	0.5854656	1.466821
3	1.394737	-0.01997376	1.1177329	1.671741	0.5540079	1.466821
4	1.394737	-0.03659426	1.0197962	1.769677	0.7498813	1.466821
5	1.388889	-0.03155007	1.0407470	1.737031	0.6962837	1.466821
6	1.611111	-0.02537723	1.2988786	1.923344	0.6244651	1.466821
7	1.500000	-0.03703704	1.1227978	1.877202	0.7544044	1.466821
8	1.617647	-0.04579687	1.1982031	2.037091	0.8388879	1.466821
9	1.550000	-0.02762500	1.2242329	1.875767	0.6515342	1.466821
10	1.250000	-0.03062500	0.9070000	1.593000	0.6860000	1.466821
11	1.550000	-0.02762500	1.2242329	1.875767	0.6515342	1.466821
12	1.447368	-0.02784662	1.1202972	1.774440	0.6541425	1.466821
13	1.555556	-0.02794925	1.2278822	1.883229	0.6553467	1.466821
14	1.342105	-0.02901298	1.0082546	1.675956	0.6677013	1.466821
15	1.722222	-0.02743484	1.3975783	2.046866	0.6492879	1.466821
16	1.333333	-0.03549383	0.9640731	1.702594	0.7385204	1.466821
17	1.558824	-0.02788520	1.2315258	1.886121	0.6545954	1.466821
18	1.447368	-0.02784662	1.1202972	1.774440	0.6541425	1.466821
19	1.500000	-0.01730104	1.2421945	1.757805	0.5156110	1.466821
20	1.277778	-0.02743484	0.9531338	1.602422	0.6492879	1.466821



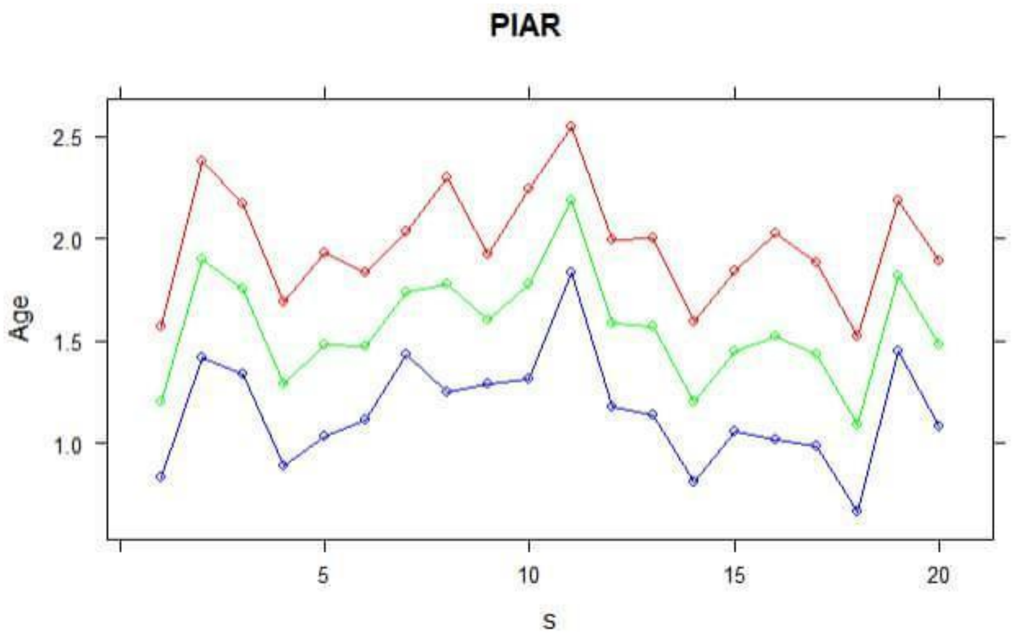
3. Sondage par PISR :

	ybarPISR	varPISR	icPISRmin	icPISRmax	delta
1	1.309250	0.07407893	-2.305254	1.842712	-4.147966
2	1.120750	0.05873098	-2.128637	1.595746	-3.724382
3	2.045062	0.15917115	-2.909946	2.827029	-5.736975
4	1.715938	0.12329061	-2.658116	2.404148	-5.062263
5	1.384125	0.11021196	-2.395964	2.034809	-4.430774
6	1.894437	0.12116886	-2.782652	2.576700	-5.359352
7	1.608875	0.13589880	-2.588962	2.331418	-4.920380
8	1.284187	0.07184514	-2.282397	1.809545	-4.091942
9	2.012063	0.12681608	-2.866481	2.710043	-5.576524
10	1.582188	0.12762362	-2.562891	2.282387	-4.845278
11	2.303688	0.19164266	-3.096136	3.161717	-6.257853
12	1.589500	0.11074774	-2.555714	2.241764	-4.797478
13	2.111875	0.15859322	-2.953342	2.892421	-5.845763
14	1.215812	0.07142186	-2.223744	1.739620	-3.963364
15	2.346437	0.18940286	-3.121167	3.199438	-6.320605
16	1.928125	0.17210011	-2.840462	2.741230	-5.581692
17	1.267562	0.08675483	-2.280953	1.844864	-4.125817
18	1.723000	0.12735077	-2.666141	2.422450	-5.088591
19	2.228625	0.17352028	-3.037776	3.045078	-6.082854
20	1.983562	0.16707990	-2.874353	2.784721	-5.659074



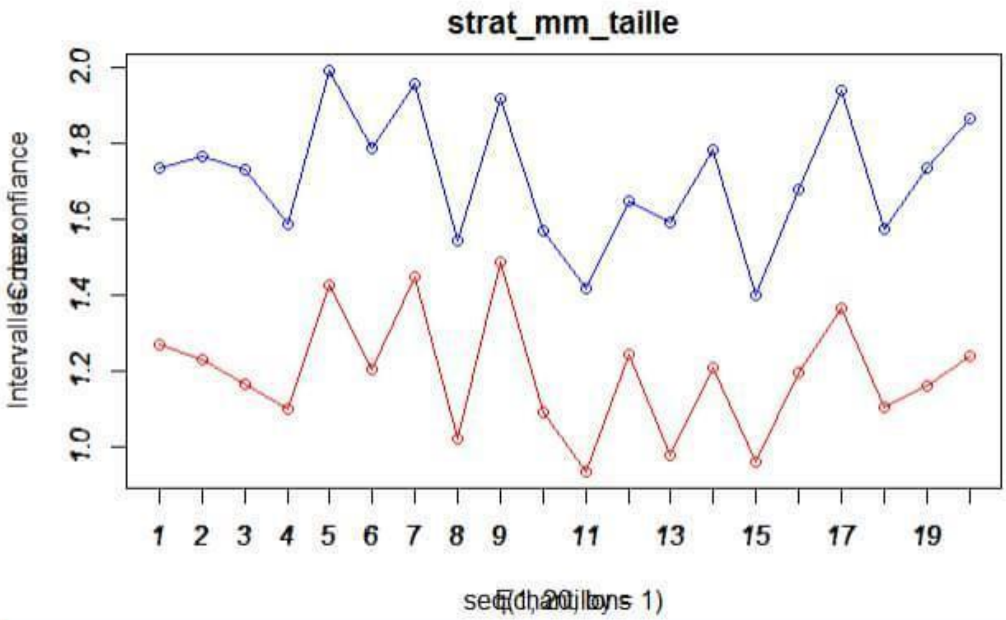
4. Sondage par PIAR :

	ybarPIAR	varPIAR	icPIARmin	icPIARmax	delta
1	1.460250	0.04523584	1.0433831	1.877117	0.8337337
2	1.513125	0.05590965	1.0496787	1.976571	0.9268927
3	1.482187	0.04801973	1.0526848	1.911690	0.8590054
4	1.861875	0.07456363	1.3266707	2.397079	1.0704086
5	1.584000	0.04989800	1.1461779	2.021822	0.8756441
6	1.688625	0.03259644	1.3347571	2.042493	0.7077358
7	1.513688	0.06324227	1.0207864	2.006589	0.9858022
8	1.438313	0.02823761	1.1089531	1.767672	0.6587188
9	1.576688	0.06440184	1.0792881	2.074087	0.9947987
10	1.648688	0.05151592	1.2038240	2.093551	0.8897270
11	1.180125	0.03325123	0.8227206	1.537529	0.7148089
12	1.681875	0.03396571	1.3206511	2.043099	0.7224477
13	1.158750	0.03166660	0.8099658	1.507534	0.6975684
14	1.461937	0.03593341	1.0903978	1.833477	0.7430795
15	1.682438	0.03225044	1.3304527	2.034422	0.7039696
16	2.143687	0.08119949	1.5851752	2.702200	1.1170245
17	1.607063	0.03815678	1.2242008	1.989924	0.7657234
18	1.488375	0.03603041	1.1163341	1.860416	0.7440818
19	1.641375	0.04339821	1.2330631	2.049687	0.8166237
20	1.747125	0.07816107	1.1991619	2.295088	1.0959262



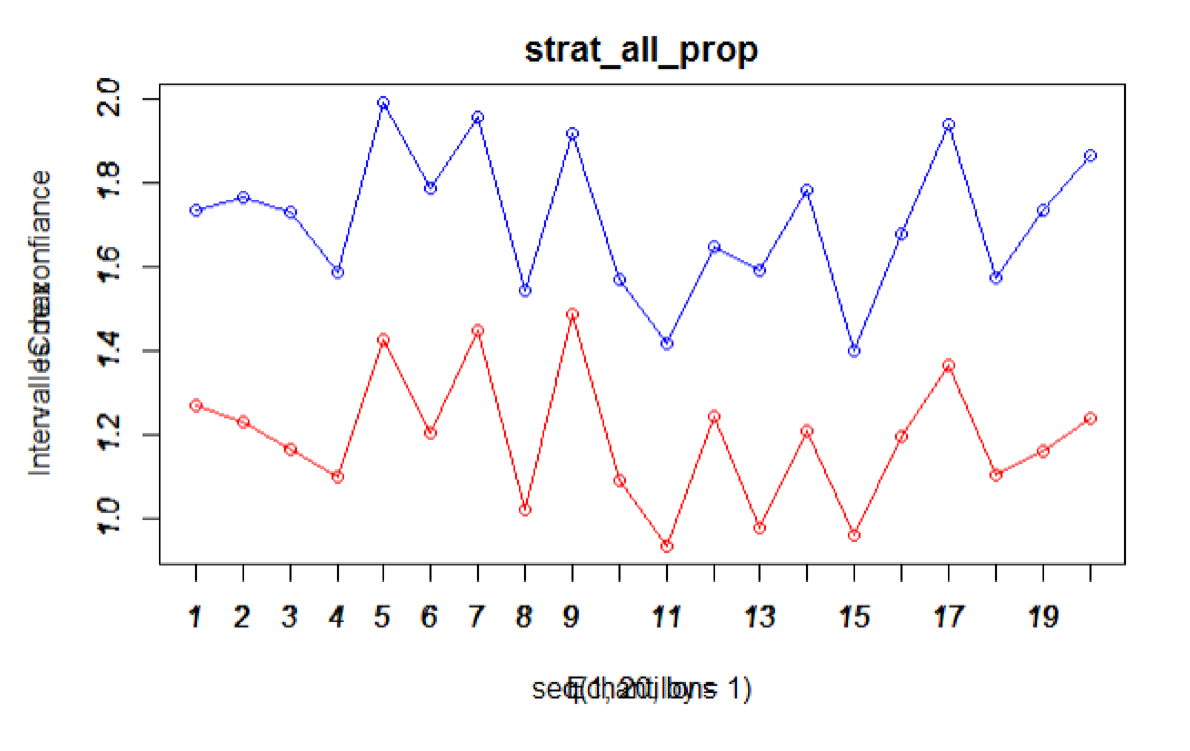
5. Sondage par strates de même taille :

	Y.bar_strat	var_Y.bar_strat	ICmin	ICmax	longueur_IC	var.inter	var.intra	moy_strat
1	1.288	0.0152500628930818	1.0562222305742	1.5197777694258	0.463555538851603	0.02074048	0.388444444444444	1.4168
2	1.38	0.0152500628930818	1.05221673502145	1.70778326497855	0.655566529957105	0.050048	0.776888888888889	1.4168
3	1.38	0.0152500628930818	1.09119414964374	1.66880585035626	0.577611700712512	0.022448	0.603111111111111	1.4168
4	1.426	0.0152500628930818	1.13842050698981	1.71357949301019	0.575158986020387	0.07164592	0.598	1.4168
5	1.334	0.0152500628930818	1.10070236006337	1.56729763993663	0.466595279873254	0.06987952	0.393555555555556	1.4168
6	1.564	0.0152500628930818	1.32620102607454	1.80179897392546	0.475597947850913	0.16421632	0.408888888888889	1.4168
7	1.242	0.0152500628930818	0.996882428210461	1.48711757178954	0.490235143579079	0.01303088	0.434444444444444	1.4168
8	1.334	0.0152500628930818	1.11965090156476	1.54834909843524	0.428698196870479	0.03307952	0.332222222222222	1.4168
9	1.38	0.0152500628930818	1.15440408514337	1.60559591485663	0.451191829713261	0.050048	0.368	1.4168
10	1.38	0.0152500628930818	1.14519232721225	1.61480767278775	0.469615345575504	0.022448	0.398666666666667	1.4168
11	1.518	0.0152500628930818	1.27871939819534	1.75728060180466	0.478561203609319	0.03673008	0.414	1.4168
12	1.518	0.0152500628930818	1.27871939819534	1.75728060180466	0.478561203609319	0.03673008	0.414	1.4168
13	1.518	0.0152500628930818	1.22554593112764	1.81045406887236	0.584908137744723	0.03673008	0.618444444444444	1.4168
14	1.38	0.0152500628930818	1.1168047660006	1.6431952339994	0.526390467998804	0.022448	0.500888888888889	1.4168
15	1.61	0.0152500628930818	1.37670236006337	1.84329763993663	0.466595279873254	0.020332	0.393555555555556	1.4168
16	1.518	0.0152500628930818	1.24556705045094	1.79043294954906	0.544865899098117	0.0183300799999999	0.536666666666667	1.4168
17	1.196	0.0152500628930818	0.983306131729192	1.40869386827081	0.425387736541617	0.04675072	0.327111111111111	1.4168
18	1.61	0.0152500628930818	1.33756705045094	1.88243294954906	0.544865899098117	0.075532	0.536666666666667	1.4168
19	1.38	0.0152500628930818	1.14519232721225	1.61480767278775	0.469615345575504	0.022448	0.398666666666667	1.4168
20	1.38	0.0152500628930818	1.11148715337995	1.64851284662005	0.537025693240091	0.096048	0.521333333333333	1.4168



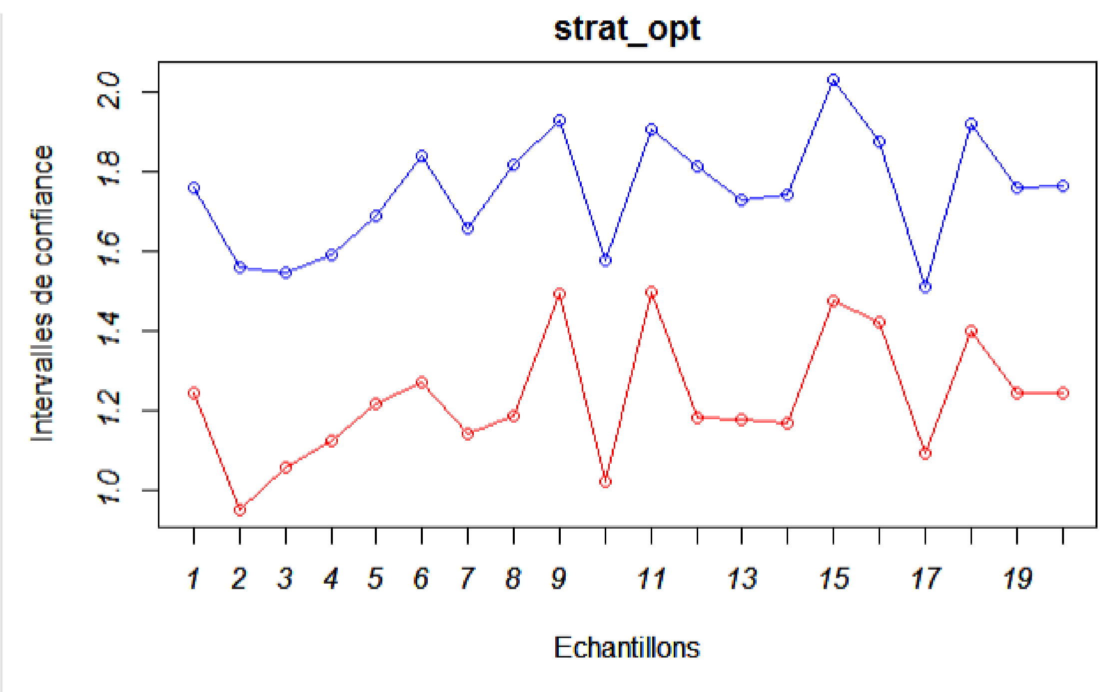
6. Sondage par strates à allocation proportionnelle :

	Y.bar_prop	var_Y.bar_prop	ICmin	ICmax	longueur_IC	var.inter	var.intra
1	1.49711111111111	0.0185232704402516	1.26643904435887	1.73356095564113	0.467121911282269	0.0110706666666667	0.225111111111111
2	1.38622222222222	0.0185232704402516	1.23121166812979	1.76301055409243	0.531798885962646	0.0961706666666667	0.468444444444444
3	1.33222222222222	0.0185232704402516	1.16223965040901	1.72976034959099	0.567520699181977	0.129566666666667	0.239444444444444
4	1.76133333333333	0.0185232704402516	1.09857355347287	1.58320422430491	0.484630670832038	0.004416	0.331
5	1.65911111111111	0.0185232704402516	1.42357298374235	1.99109368292432	0.567520699181977	0.00196266666666666	0.382111111111111
6	1.32933333333333	0.0185232704402516	1.20471371611038	1.78373072833407	0.579017012223686	0.216384	0.379
7	1.44311111111111	0.0185232704402516	1.44697342648039	1.95613768463072	0.509164258150332	0.0240426666666667	0.382111111111111
8	1.70733333333333	0.0185232704402516	1.0222393710973	1.53998285112492	0.517743480027621	0.013524	0.524
9	1.70733333333333	0.0185232704402516	1.48621308309728	1.91689802801383	0.430684944916557	0.013524	0.529
10	1.65333333333333	0.0185232704402516	1.08960361886306	1.56906304780361	0.479459428940552	0.0276	0.475
11	1.60511111111111	0.0185232704402516	0.934673134524977	1.41732686547502	0.482653730950047	3.06666666666665e-05	0.460111111111111
12	1.76133333333333	0.0185232704402516	1.240514561837	1.64570766038522	0.405193098548224	0.004416	0.451
13	1.70155555555556	0.0185232704402516	0.978089267922814	1.58991073207719	0.611821464154372	0.106750666666667	0.536777777777778
14	1.5	0.0185232704402516	1.20737986985579	1.78106457458866	0.573684704732873	0	0.47
15	1.33511111111111	0.0185232704402516	0.958491752388031	1.39928602538975	0.440794273001716	0.0648906666666667	0.310111111111111
16	1.59355555555556	0.0185232704402516	1.19600646785831	1.67866019880836	0.482653730950047	0.181822666666667	0.488777777777778
17	1.34377777777778	0.0185232704402516	1.36191066748853	1.93897822140036	0.577067553911826	0.00371066666666667	0.498444444444444
18	1.33511111111111	0.0185232704402516	1.10247355987066	1.57352644012934	0.471052880258682	0.0648906666666667	0.430111111111111
19	1.38911111111111	0.0185232704402516	1.16088904615922	1.73111095384078	0.570221907681562	0.0419826666666667	0.412111111111111
20	1.34088888888889	0.0185232704402516	1.23723900693844	1.86498321528378	0.627744208345335	0.00196266666666666	0.382111111111111



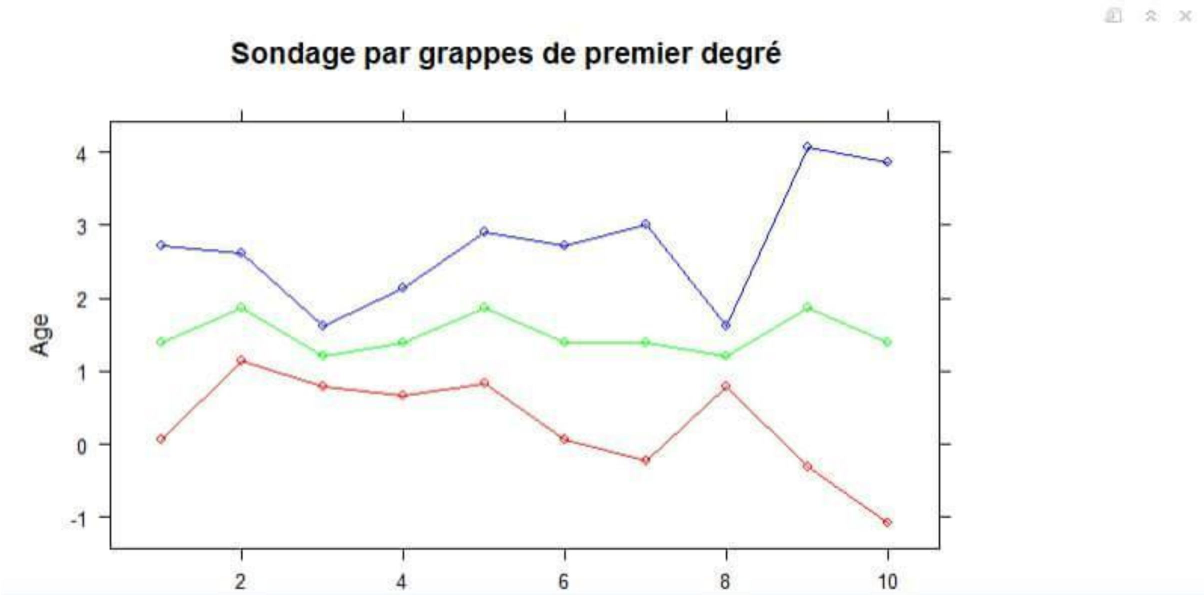
7. Sondage par strates à allocation optimale :

	y.bar_opt	var_Y.bar_opt	ICmin	ICmax	longueur_IC	var.inter	var.intra
1	1.5	0.0184995934184613	1.24150852762049	1.75849147237951	0.516982944759029	0	0.439
2	1.25454545454545	0.0184995934184613	0.9498793807747	1.55921152831621	0.60933214754151	0.0513223140495868	0.600272727272727
3	1.30161616161616	0.0184995934184613	1.05636971472761	1.54686260850471	0.490492893777099	0.00648815426997245	0.391929292929293
4	1.35676767676768	0.0184995934184613	1.12351974577904	1.59001560775631	0.466495861977268	0.113770798898072	0.351717171717172
5	1.45292929292929	0.0184995934184613	1.21687298095649	1.6889856049021	0.472112623945608	0.0213146005509642	0.362474747474748
6	1.55515151515152	0.0184995934184613	1.26874053850338	1.84156249179965	0.572821953296262	0.0659206611570248	0.534363636363636
7	1.39979797979798	0.0184995934184613	1.14280873988614	1.65678721970982	0.513978479823684	0.000101377410468321	0.431202020202
8	1.5020202020202	0.0184995934184613	1.18512236790193	1.81891803613847	0.633795668236539	0.010137741046832	0.6542020202020202
9	1.70848484848485	0.0184995934184613	1.49019495017035	1.92677474679934	0.436579796628989	0.178829752066116	0.311363636363636
10	1.29959595959596	0.0184995934184613	1.02031975905873	1.57887216013319	0.558552401074462	0.000405509641873279	0.508808080808081
11	1.70242424242424	0.0184995934184613	1.49808915002173	1.90675933482676	0.408670184805029	0.014598347107438	0.272090909090909
12	1.4979797979798	0.0184995934184613	1.18108196386153	1.81487763209807	0.633795668236539	0.010137741046832	0.6542020202020202
13	1.45090909090909	0.0184995934184613	1.17498591713507	1.72683226468311	0.55184634754804	0.00205289256198347	0.495090909090909
14	1.4549494949495	0.0184995934184613	1.16786338682901	1.74203560306998	0.574172216240973	0.0608517906336088	0.538262626262626
15	1.75353535353535	0.0184995934184613	1.4739792348319	2.03309147223881	0.559112237406911	0.0310468319559229	0.507868686868687
16	1.64727272727273	0.0184995934184613	1.41907606429229	1.87546939025317	0.45639332596088	0.0184760330578512	0.340818181818182
17	1.30161616161616	0.0184995934184613	1.09202779604167	1.51120452719065	0.419176731148976	0.00648815426997245	0.283929292929293
18	1.65939393939394	0.0184995934184613	1.39806607305994	1.92072180572794	0.522655732668002	0.219203305785124	0.448818181818182
19	1.5020202020202	0.0184995934184613	1.24503096210836	1.75900944193204	0.513978479823684	0.010137741046832	0.431202020202
20	1.5020202020202	0.0184995934184613	1.24127671923625	1.76276368480415	0.521486965567902	0.010137741046832	0.438202020202



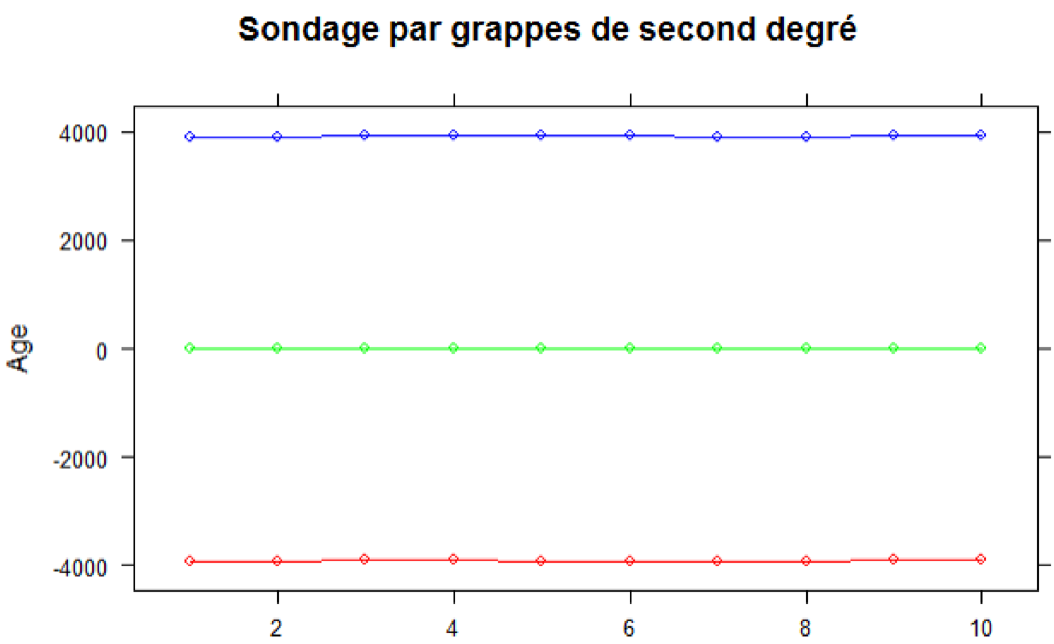
8. Sondage par grappes de 1er degré :

	ybar_chap	var_estim_ybar_Grappe	icSgmin	icSgmax	delta
1	1.875	0.6783652	0.5454042	3.204596	2.6591916
2	1.875	0.3768695	1.1363358	2.613664	1.4773284
3	1.875	0.2119891	1.4595014	2.290499	0.8309973
4	1.875	0.3768695	1.1363358	2.613664	1.4773284
5	1.395	0.5299728	0.3562533	2.433747	2.0774934
6	1.200	0.6783652	-0.1295958	2.529596	2.6591916
7	1.875	0.8244021	0.2591719	3.490828	3.2316562
8	1.875	0.2119891	1.4595014	2.290499	0.8309973
9	1.395	1.1129430	-0.7863683	3.576368	4.3627366
10	1.875	1.2562320	-0.5872147	4.337215	4.9244294



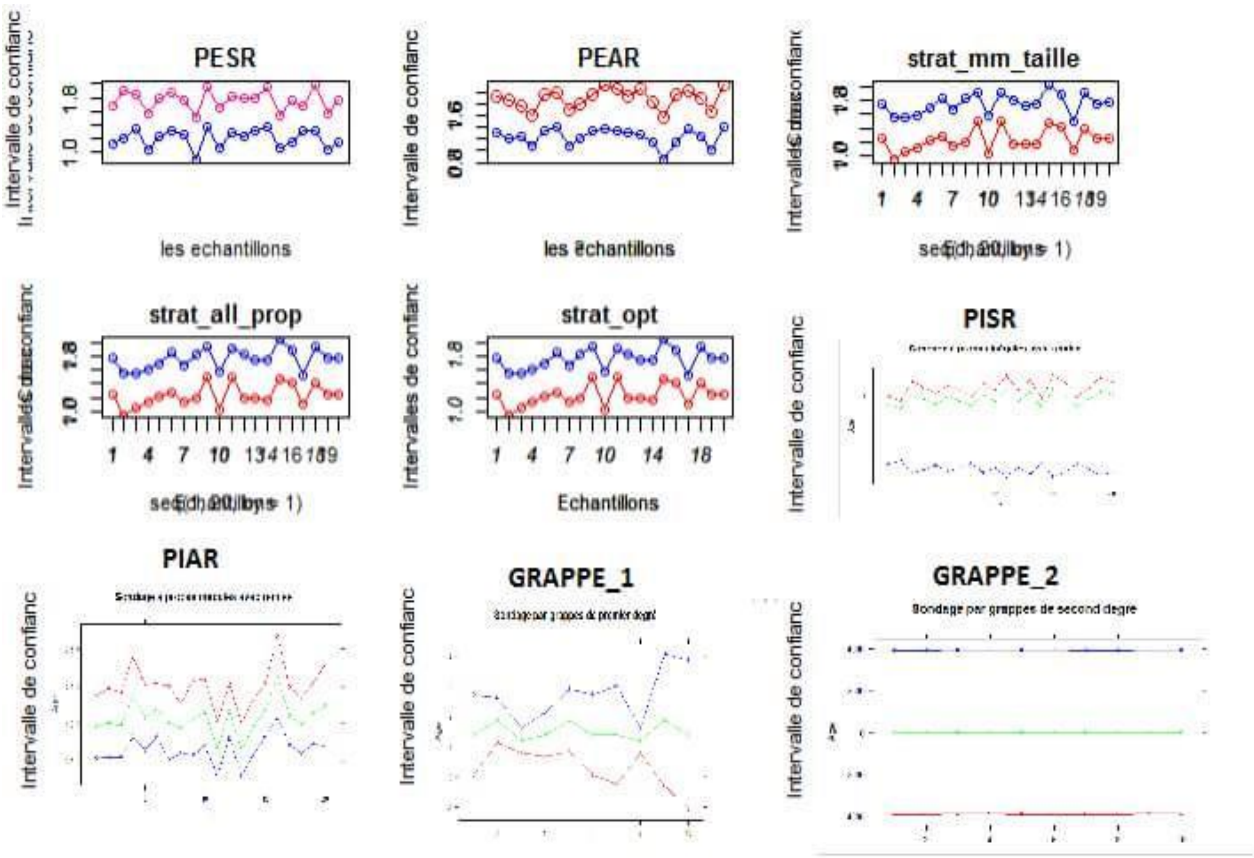
9. Sondage par grappe de 2nd degré :

	ybar_chap	var_estim_ybar_Grappe	icSgmin	icSgmax	delta
1	0.36	39857.35	-390.9404	391.6604	782.6008
2	0.20	39857.35	-391.1004	391.5004	782.6008
3	1.20	39857.35	-390.1004	392.5004	782.6008
4	1.32	39857.35	-389.9804	392.6204	782.6008
5	1.08	39857.35	-390.2204	392.3804	782.6008
6	0.96	39857.35	-390.3404	392.2604	782.6008
7	0.47	39857.35	-390.8304	391.7704	782.6008
8	0.36	39857.35	-390.9404	391.6604	782.6008
9	1.44	39857.35	-389.8604	392.7404	782.6008
10	1.56	39857.35	-389.7404	392.8604	782.6008



Conclusion :

Nous avons utilisé différentes méthodes de sondage dans ce travail et nous avons abouti à des résultats différents.
Dans chaque méthode on a estimé la variable d'intérêt "moy_eau " par sa moyenne estimée et on a tracé son Intervalle de confiance ainsi que la longueur de cet Intervalle.
Finalement on constate que la meilleure méthode de sondage est le sondage aléatoire simple sans remise, dont les résultats sont plus précis.



Annexes : code R :

Statistique descriptive :

Le sexe :

```
freq=table(data$sexe)/100
x=c(freq[1],freq[2],freq[3])
names(x)=c("Femme", "Homme")
pie(freq,labels=names(x),main="Répartition selon le sexe",col=c("red","light blue"))
```

L'âge :

```
freq=table(data$age)/100
x=c(1,2,3,4)
names(x)=c("Moins de 18 ans","Entre 18 et 25 ans","Entre 25 et 35 ans","35 ans et plus")
pie(freq,label=names(x),main="Age de la population")
```

Consommation en eau :

```
freq=table("litre eau")/100
x=c(freq[1],freq[2],freq[3])
names(x)=c("Moins d'un litre", "1 a 2 litres","Plus de 2 litres")
pie(freq,labels=names(x),main="litre eau ")
```

Maladie :

```
freq=table("maladie contrainte")/100
x=c(1,2,3,4)
names(x)=c("Oui","Non")
pie(freq,label=names(x),col=rainbow(4),main="Souffrez-vous d'une maladie ")
```

Activité physique :

```
freq=table("Activite physique" )/100
x=c(freq[1],freq[2])
names(x)=c("Non", "Oui")
pie(freq,labels=names(x),main="Activité physique")
```

Sondage :

```
library(MASS)
library(sampling)
library(lpSolve)
```

PESR

```
N=100
n=20
t=srswor(n,N)

ybarPESR=NULL
varPESR=NULL


#Selection de l echantillon
for (i in 1:20) {
  t=srswor(n,N)
  echPESR=as.vector(data$moy_eau[t==1])
 ybarPESR[i]=mean(echPESR)
 s2y=sum((echPESR-ybarPESR[i])^2)/(n-1)
 varPESR[i]=(1-n/N)*s2y*1/n }


icPESRmin=c(-1.96*sqrt(varPESR)+ybarPESR)  # Borne inferieure de l IC
icPESRmax=c(1.96*sqrt(varPESR)+ybarPESR)  # Borne superieure de l IC
longueurIC=icPESRmax-icPESRmin  # Longueur des intervalles de confiance
moy_PESR=mean(ybarPESR)

plot(icPESRmin,ylim=c(min(min(icPESRmax),min(icPESRmin)),max(max(icPESRmax),max(icPESRmin))),type="o",xaxt ="n",xlab="les echantillons",ylab="Intervalle de confiance",col="blue",main="PESR")

par(new=T)
plot(icPESRmax,ylim=c(min(min(icPESRmax),min(icPESRmin)),max(
max(icPESRmax),max(icPESRmin))),type="o",col="deeppink",xaxt =
"n",xlab="",ylab="",main="PESR")
```

```
PESR=cbind(ybarPESR,varPESR,icPESRmin,icPESRmax,longueurIC,moy_PESR)
View(PESR)
```

PEAR

```
ybarPEAR=NULL
varPEAR=NULL
for (i in 1:20) {t=srswr(n,N) ## Seelection de l echantillon
  echPEAR=data$moy_eau[t!=0]
  m=length(echPEAR) ## Taille de l echantillon tire
  ybarPEAR[i]=mean(echPEAR)
  s2y=sum((echPEAR-ybarPEAR[i])^2)/m-1
  varPEAR[i]=s2y/m }
icPEARmin=c(-1.96*sqrt(abs(varPEAR))+ybarPEAR) # Borne inferieure de l IC
icPEARmax=c(1.96*sqrt(abs(varPEAR))+ybarPEAR) # Borne superieure de l IC
longueurIC=icPEARmax-icPEARmin # Longueur des intervalles de confiance
moy_PEAR=mean(ybarPEAR)
plot(icPEARmin,ylim=c(min(min(icPEARmax),min(icPEARmin)),max(max(icPEARmax),max(icPEARm
in))),type="o",xaxt
     = "n",xlab="les ?chantillons",ylab="Intervalle de confiance",col="blue")

par(new=T)
plot(icPEARmax,ylim=c(min(min(icPEARmax),min(icPEARmin)),max(max(icPEARmax),max(icPEAR
min))),type="o",xax= "n",xlab="les echantillons",ylab="Intervalle de
confiance",cex=1.5,font=3,col="red")

PEAR=cbind(ybarPEAR,varPEAR,icPEARmin,icPEARmax,longueurIC,moy_PEAR)
View(PEAR)
```

PISR

```
y=data$moy_eau  #notre variable d'intérêt sera l'âge ou on commence à fumer

x=data$moy_age  # Nous avons choisi l’âge comme variable

ybarPISR=NULL

varPISR=NULL

icPISRmin=NULL

icPISRmax=NULL

for(i in 1:20)

  {
```

```
 pi=inclusionprobabilities(x,20)

 pi

 sp=UPpoisson(pi)  # Nous avons utilisé la méthode de poisson pour l'échantillonnage

 sp

 ech=getdata(y,sp)

 View(ech)

 pk=getdata(pi,sp)

 View(pk)

 tybarPISR=sum(ech/pk)

 tybarPISR

 ybarPISR[i]=1/100*tybarPISR

 ybarPISR

 varPISR[i]=1/10000*sum((1-pk[,2]) *(ech[,2]/pk[,2])^2)

 varPISR

  icPISRmin[i]=-1.96*sqrt(varPISR[i]+ybarPISR[i])  #borne inff de l'intervalle de confiance

icPISRmin

  icPISRmax[i]=1.96*sqrt(varPISR[i])+ybarPISR[i]

icPISRmax

 }

delta=icPISRmin-icPISRmax  #amplitude de l intervalle de confiance

PISR_tab=data.frame(ybarPISR,varPISR,icPISRmin,icPISRmax,delta)

s=1:20

h=PISR_tab$ybarPISR

h

h1=PISR_tab$icPISRmin

h2=PISR_tab$icPISRmax

library(lattice)
```

```
PISR=xyplot(h1+h2+h ~ s,col=c("blue","red","green"),main="Sondage à probas inégales avec remise",ylab="Age",type='o')  #traçage du graphique
PISR
mean(PISR$ybarPISR)
View(PISR_tab)
```

PIAR

```
y=moy_eau # notre variable d'intérêt sera l'âge ou on commence à fumer
x=moy_age   # Nous avons choisi l'âge comme variable auxiliaire  pi=inclusionprobabilities(x,20)
pi=inclusionprobabilities(x,20)
p=pi/20
v=cumsum(p)
v
ybarPIAR=NULL
varPIAR=NULL
k=NULL
for(l in 1:20)
 {
 u=runif(20,min=0,max=1)
 u
 for(i in 1:20)
   {
   for (j in 1:99)
     {if ((u[i] < v[j+1])&&(u[i] >= v[j])) {k[i]=j}}
   }
 ech=y[k]
 t=1/20*sum(ech/p[k])
 t
```

```
 ybarPIAR[l]=1/100*t

 ybarPIAR

 varPIAR[l]=(1/20)*(1/19)*(1/10000)*sum((ech/p[k]- 100*ybarPIAR[l])^2)

 }

icPIARmin=-1.96*sqrt(varPIAR)+ybarPIAR #borne inf de l'intervalle de confiance

icPIARmax=1.96*sqrt(varPIAR)+ybarPIAR #borne sup de l'intervalle de confiance

delta=icPIARmax-icPIARmin

PIAR=data.frame(ybarPIAR,varPIAR,icPIARmin,icPIARmax,delta)

View(PIAR)

s=1:20

h=PIAR$ybarPIAR

h1=PIAR$icPIARmin

h2=PIAR$icPIARmax

library(lattice)

PIAR_IC=xyplot(h1+h2+h ~ s,col=c("blue","red","green"),main="Sondage à probas inégales avec remise",ylab="Age",type='o')

PIAR_IC
```

Strates de même taille

```
N=100
attach (data)
View(data)
Nh=min(table(data$sexe))
Nf=min(table(data$sexe))

nh=10
nf=10

N1=sum(data$sexe=="Homme")
N1
```

```
N2=sum(data$sexe=="Femme")
N2
ech=matrix(0,nrow=20,ncol=20)
IC=matrix(0,nrow=20,ncol=20)
y.bar.strat1=NULL
var.strat1=NULL
```

Echantillonnage

```
library(tibble)
Ech=list(NULL) #variable qui va contenir les 20 échantillons(initialisation) construction de chaque

library(sampling)

for (i in 1:20){     # échantillon selon un tirage sans remise affectation des 20 ech selon les 2
strates sélection de 20 échantillons homme et 20 echantillons femme

s=strata(data,stratanames="sexe",size=c(nf,nh),method=c("srswor"))

Ech[[i]]=getdata(data["moy_eau"],s[c("ID_unit","sexe")])
}
```

```
Ech_h=list(NULL)
Ech_f=list(NULL)
for (i in 1:20){
Ech_h[[i]]=Ech[[i]][Ech[[i]]$genre=="Homme",]
Ech_f[[i]]=Ech[[i]][Ech[[i]]$genre=="Femme",]
}
```

sélection de 20 échantillons homme et 20 echantillons femme

```
Ech_h=list(NULL)
Ech_f=list(NULL)
for (i in 1:20){
Ech_h[[i]]=Ech[[i]][Ech[[i]]$sexe=="Homme",]
Ech_f[[i]]=Ech[[i]][Ech[[i]]$sexe=="Femme",]
}
```

calculs

Estimation de la moyenne sur chaque échantillon

```
ypesr.hat_f=list(NULL)
ypesr.hat_h=list(NULL)
for (i in 1:20){
```

```
ypesr.hat_h[[i]]=mean(Ech_h[[i]][,1])
ypesr.hat_f[[i]]=mean(Ech_f[[i]][,1])
}
```

Estimation de la moyenne de la population

```
ystrat.hat=list(NULL)
for (i in 1:20){
ystrat.hat[[i]]=(Nf/N)*ypesr.hat_f[[i]]+(Nh/N)*ypesr.hat_h[[i]]
}

ystrat.hat
```

calcul de la variation de l'estimateur de la moyenne

```
fh=nh/Nh  #taux de sondage pour la strate Homme
ff=nf/Nf  # taux de sondage pour la strate Femme
strat_h=data[data$sexe=="Homme",] # strat_h
strat_hstrat_f=data[data$sexe=="Femme",]  “strat_f
strat_f

var_ystrat.hat=(1/N^2)*sum((Nh^2)*((1-fh)/nh)*var(strat_h[,"moy_eau"]),
(Nf^2)*((1-ff)/nf)*var(strat_f[,"moy_eau"]))
var_ystrat.hat
```

calcul des intervalles de confiance

```
#calcul des intervalles de confiance
for (i in 1:20){
s2c_h=list(NULL)
s2c_f=list(NULL)

ICmin=list(NULL)
ICmax=list(NULL)

var_ystrat.hat_hat=list(NULL)

long=list(NULL)}
```

etablissement des calculs pour 20 échantillons

```
for (i in 1:20){
s2c_h[[i]]=var(Ech_h[[i]][,"moy_eau"])

s2c_f[[i]]=var(Ech_f[[i]][,"moy_eau"])

var_ystrat.hat_hat[[i]]=(1/N^2)*sum((Nh^2)*((1-fh)/nh)*s2c_h[[i]],(Nf^2)*((1-ff)/nf)*s2c_f[[i]])

ICmin[[i]]=ystrat.hat[[i]]-1.96*sqrt(var_ystrat.hat_hat[[i]])

ICmax[[i]]=ystrat.hat[[i]]+1.96*sqrt(var_ystrat.hat_hat[[i]])
}
```

etablissement des calculs pour 20 échantillons

```
for (i in 1:20){

s2c_h[[i]]=var(Ech_h[[i]][,"moy_eau"])
s2c_f[[i]]=var(Ech_f[[i]][,"moy_eau"])
var_ystrat.hat_hat[[i]]=(1/N^2)*sum((Nh^2)*((1-fh)/nh)*s2c_h[[i]]
,(Nf^2)*((1-ff)/nf)*s2c_f[[i]])
ICmin[[i]]=ystrat.hat[[i]]-1.96*sqrt(var_ystrat.hat_hat[[i]])
ICmax[[i]]=ystrat.hat[[i]]+1.96*sqrt(var_ystrat.hat_hat[[i]])
long[[i]]=ICmax[[i]]-ICmin[[i]]
}
s2c_h
```

Representation des IC

```
p=plot(seq(1,20, by =1),ICmin,ylim=c(min(min(as.numeric(ICmax)),min(as.numeric(ICmin)))
, max(max(as.numeric(ICmax)),max(as.numeric(ICmin)))), type="o", xaxp=c(0,20,20) , col="red"
,xlab="Echantillons",ylab="Intervalles de confiance" , cex=1 , font=3)
p

par(new=TRUE)

p1=plot(seq(1,20, by =1),ICmax,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)),
max(as.numeric(ICmin)))),type="o",xaxp=c(0,20,20),col="blue")
```

calcul des variances inter-strates et intra-strate

```
var.inter=list(NULL)
var.intra=list(NULL)
```

```
for(i in 1:20){
var.inter[[i]]=(1/N)*sum(Nh*(ypesr.hat_h[[i]]-ystrat.hat[[i]])^2,Nf*
(ypesr.hat_f[[i]]-ystrat.hat[[i]])^2)
var.intra[[i]]=(1/N)*sum(Nh*var(Ech_h[[i]][,"moy_eau"]),
Nf*var(Ech_f[[i]][,"moy_eau"]))
}
var.inter
```

moyenne des estimateurs

```
moy_strat=mean(as.numeric(ystrat.hat))
res1=cbind(ystrat.hat,var_ystrat.hat,ICmin,ICmax,long,var.inter,var.intra,moy_strat)

colnames(res1)=c("Y.bar_strat","var_Y.bar_strat","ICmin","ICmax","longueur_IC",
"var.inter","var.intra","moy_strat")
View(res1)
res1
```

Strates à allocation proportionnelle

```
N=100
n=20
Nf=table(data$sexe)[1]
Nh=table(data$sexe)[2]
nf=(n/N)*Nf
nh=(n/N)*Nh
```

initialisation de 20 échantillons

```
Ech=list(NULL)
for (i in 1:20){

str=strata(data,stratanames="sexe",size=c(nf,nh),method=c("srswor"))

Ech[[i]]=getdata(data["moy_eau"],str[c("ID_unit","sexe")])
}
```

sélection de 20 échantillons homme et 20 echantillons femme

```
Ech_h=list(NULL)
Ech_f=list(NULL)
for (i in 1:20){
Ech_h[[i]]=Ech[[i]][Ech[[i]]$sexe=="Homme",]
Ech_f[[i]]=Ech[[i]][Ech[[i]]$sexe=="Femme",]
}
```

calculs

Estimation de la moyenne sur chaque échantillon

```
ypesr.hat_f=list(NULL)
ypesr.hat_h=list(NULL)
for (i in 1:20){
ypesr.hat_h[[i]]=mean(Ech_h[[i]][,1])
ypesr.hat_f[[i]]=mean(Ech_f[[i]][,1])
}
```

Estimation de la moyenne de la population

```
yprop.hat=list(NULL)
for (i in 1:20){
yprop.hat[[i]]=(Nf/N)*ypesr.hat_f[[i]]+(Nh/N)*ypesr.hat_h[[i]]
}
```

calcul de la variance de l'estimateur de la moyenne

```
fh=nh/Nh
ff=nf/Nf
strat_h=data[data$sexe=="Homme",]
strat_f=data[data$sexe=="Femme",]
var_yprop.hat=(1/N^2)*sum((Nh^2)*((1-fh)/nh)*var(strat_h[,"moy_eau"])
,(Nf^2)*((1-ff)/nf)*var(strat_f[,"moy_eau"]))
```

calcul des intervalles de confiance

```
s2c_h=list(NULL)
s2c_f=list(NULL)

ICmin=list(NULL)
ICmax=list(NULL)

var_yprop.hat_hat=list(NULL)
```

```
long=list(NULL)
```

etablissement des calculs pour 20 échantillons

```
for (i in 1:20){

s2c_h[[i]]=var(Ech_h[[i]][,"moy_eau"])

s2c_f[[i]]=var(Ech_f[[i]][,"moy_eau"])

var_yprop.hat_hat[[i]]=(1/N^2)*sum((Nh^2)*((1-fh)/nh)*s2c_h[[i]],(Nf^2)*((1-ff)/nf)*s2c_f[[i]])

ICmin[[i]]=yprop.hat[[i]]-1.96*sqrt(var_yprop.hat_hat[[i]])

ICmax[[i]]=yprop.hat[[i]]+1.96*sqrt(var_yprop.hat_hat[[i]])

long[[i]]=ICmax[[i]]-ICmin[[i]]
}
```

Representation des IC

```
p1=plot(seq(1,20, by =1),ICmin,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),
max(max(as.numeric(ICmax)),
max(as.numeric(ICmin)))),type="o",xaxp=c(0,20,20),
col="red",xlab="Echantillons",ylab="Intervalles de confiance",cex=1,font=3)
p1


par(new=TRUE)
plot(seq(1,20, by =1),ICmax,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)),
max(as.numeric(ICmin)))),type="o",xaxp=c(0,20,20),col="blue")
```

calcul des variances interstrates et intra-strate(respectivement)

```
var.inter=list(NULL)
var.intra=list(NULL)
for(i in 1:20) {
var.inter[[i]]=(1/N)*sum(Nh*(ypesr.hat_h[[i]]-yprop.hat[[i]])^2,Nf* (ypesr.hat_f[[i]]-yprop.hat[[i]])^2)
```

```
var.intra[[i]]=(1/N)*sum(Nh*var(Ech_h[[i]][,"moy_eau"]),Nf*var(Ech_f[[i]][,"moy_eau"]))
}
```

moyenne des estimateurs de la moyenne

```
moy_prop=mean(as.numeric(yprop.hat))

res2=cbind(yprop.hat,var_yprop.hat,ICmin,ICmax,long,var.inter,var.intra)

colnames(res2)=c("Y.bar_prop","var_Y.bar_prop","ICmin","ICmax","longueur_IC","var.inter","var.intra"
)

View(res2)
```

Strates à allocation optimale :

```
N=100
n=20
strat_h=data[data$sexe=="Homme",]
strat_f=data[data$sexe=="Femme",]
sigma2c_h= var(strat_h[,"moy_eau"])  #variance corrigee sur chaque strate
sigma2c_f= var(strat_f[,"moy_eau"])

nh=round(n*(Nh*sqrt(sigma2c_h))/sum(Nh*sqrt(sigma2c_h),Nf*sqrt(sigma2c_f)))

nf=round(n*(Nf*sqrt(sigma2c_f))/sum(Nh*sqrt(sigma2c_h),Nf*sqrt(sigma2c_f)))
```

Echantillonnage

```
Ech=list(NULL)

for (i in 1:20){

str=strata(data,stratanames="sexe",size=c(nf,nh),method=c("srswor"))

Ech[[i]]=getdata(data["moy_eau"],str[c("ID_unit","sexe")])
}
```

sélection de 20 échantillons homme et 20 echantillons femme

```
Ech_h=list(NULL)
Ech_f=list(NULL)
for (i in 1:20){
  Ech_h[[i]]=Ech[[i]][Ech[[i]]$sexe=="Homme",]
  Ech_f[[i]]=Ech[[i]][Ech[[i]]$sexe=="Femme",]
  }
```

calculs

Estimation de la moyenne sur chaque échantillon

```
ypesr.hat_f=list(NULL)
ypesr.hat_h=list(NULL)
for (i in 1:20) {
  ypesr.hat_h[[i]]=mean(Ech_h[[i]][,1])
  ypesr.hat_f[[i]]=mean(Ech_f[[i]][,1])
  }
```

Estimation de la moyenne de la population

```
yopt.hat=list(NULL)
for (i in 1:20) {
  yopt.hat[[i]]=(Nf/N)*ypesr.hat_f[[i]]+(Nh/N)*ypesr.hat_h[[i]]
  }
```

calcul de la variance de l'estimateur de la moyenne

```
fh=nh/Nh
ff=nf/Nf
var_yopt.hat=(1/N^2)*sum((Nh^2)*((1-fh)/nh)*var(strat_h[,"moy_eau"]),(Nf^2)*((1-
ff)/nf)*var(strat_f[,"moy_eau"]))
```

variance corrigée sur chaque échantillon

```
s2c_h=list(NULL)
s2c_f=list(NULL)

ICmin=list(NULL)
```

```
ICmax=list(NULL)

var_yopt.hat_hat=list(NULL) #variance estimée

long=list(NULL) # longueur de chaque IC
```

etablissement des calculs pour 20 échantillons

```
for (i in 1:20) {
s2c_h[[i]]=var(Ech_h[[i]][,"moy_eau"])  #variance empirique corrigee sur chaque echantillon

s2c_f[[i]]=var(Ech_f[[i]][,"moy_eau"])

var_yopt.hat_hat[[i]]=(1/N^2)*sum((Nh^2)*((1-fh)/nh)*s2c_h[[i]], (Nf^2)*((1-ff)/nf)*s2c_f[[i]]) #variance corrigee sur lechantillon final

ICmin[[i]]=yopt.hat[[i]]-1.96*sqrt (var_yopt.hat_hat[[i]])  #borne inferieure

ICmax[[i]]=yopt.hat[[i]]+1.96*sqrt (var_yopt.hat_hat[[i]])

long[[i]]=ICmax[[i]]-ICmin[[i]]
}
```

Representation des IC

```
plot(seq(1,20, by =1),ICmin,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)), max(as.numeric(ICmin)))),
type="o",xaxp=c(0,20,20), col="red",xlab="Echantillons",ylab="Intervalles de confiance",cex=1,font=3)

par(new=TRUE)

plot(seq(1,20, by =1),ICmax,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)), max(as.numeric(ICmin)))),
type="o",xaxp=c(0,20,20), col="blue",xlab="Echantillons",ylab="Intervalles de
confiance",cex=1,font=3)
```

calcul des variances inter-strates et intra-strate(respectivement)

```
var.inter=list(NULL)
var.intra=list(NULL)
```

```
for(i in 1:20) {
var.inter[[i]]=(1/N)*sum(Nh* (ypesr.hat_h[[i]]-yopt.hat[[i]])^2,Nf* (ypesr.hat_f[[i]]-yopt.hat[[i]])^2)

var.intra[[i]]=(1/N)*sum(Nh*var(Ech_h[[i]] [,"moy_eau"]),Nf*var(Ech_f[[i]][,"moy_eau"]))
 }
```

moyenne des estimateurs de la moyenne

```
moy_opt=mean(as.numeric(yopt.hat))

res3=cbind(yopt.hat,var_yopt.hat,ICmin,ICmax,long,var.inter,var.intra)

colnames(res3)=c("y.bar_opt","var_Y.bar_opt","ICmin","ICmax","longueur_IC", "var.inter","var.intra")
View(res3)
```

Sondage par grappe 1er degré :

```
Ni=table(data$moy_age)
Ni
ybar_chap=NULL
var_estim_ybar_Grappe=NULL
icSgmin=NULL
icSgmax=NULL
for (l in 1:10) {
  Sg=cluster(data,clustername=c("sport"),size=2,method="srswor")
  pik=Sg$Prob  #Probabilité d'inclusion des individus dans chaque unité p rimaire
  Ech_Sg=getdata(data,Sg)
  Ech_Sg
  n=dim(Sg)[1]
  y_ech=Ech_Sg$moy_eau
  gr=unique(Sg$moy_age)
  ty=NULL
  for (i in 1:2)  {
  ty[i]=sum(y_ech[Ech_Sg$moy_age==gr[i]])
  }
  pi=unique(pik)
  t_chap_y_Grappe=sum(ty/pi)
  t_chap_y_Grappe
Ni_ech=as.vector(Ni[gr])
  N=sum(Ni)
  ybar_chap[l]=t_chap_y_Grappe/N
 M=3  #Nombre de grappes
   m=2  # Nous allons choisir 2 unités primaires
   differance_carre = NULL
   for (i in 1:2)       {
```

```
  differance_carre[i] =  (ty[i]-(t_chap_y_Grappe/Ni_ech[i]))^2
   }

   var_estim_ybar_Grappe[l]=(1/10000)*((M-m)*M/((M-1)*m))*sum(differance_carre)
}

icSgmin=-1.96*var_estim_ybar_Grappe+ybar_chap
icSgmax=+1.96*var_estim_ybar_Grappe+ybar_chap
delta=icSgmax-icSgmin
SG=data.frame(ybar_chap,var_estim_ybar_Grappe,icSgmin,icSgmax,delta)
SG
s=1:10
h=SG$ybar_chap
h1=SG$icSgmin
h2=SG$icSgmax
library(lattice)
xyplot(h2+h1+h ~ s,col=c("blue","red","green"),main="Sondage par grappes de premier degre",ylab="eau",xlab="",type='o')
mean(SG$ybar_chap )
```

Sondage par grappe 2^{nd} degré :

```
attach(data)

library(MASS)

 library(lpSolve)

 library(sampling)

 Ni=as.vector(table(base$X12..raisons))

ybar_chap=NULL

 var_estim_ybar_Grappe=NULL

icSgmin=NULL

 icSgmax=NULL
```

```
 M=3

m=2 for(l in 1:10){

Sg=cluster(data,clustername=c("sport"),size=2,method="srswor")

pik=Sg$Prob

 Ech_Sg=getdata(data,Sg)

 n=dim(Sg)[1]

y_ech=Ech_Sg$moy_eau

gr=unique(Sg$sport)

 ty=c(0,0)

 nh=round(Ni/5)

 for (i in 1:2)    {

p2=srswor(nh[gr[i]],Ni[gr[i]])

     grappe=Ech_Sg$moy_eau[Ech_Sg$sport==gr[i]]
ech2=getdata(Ech_Sg$moy_eau[Ech_Sg$sport==gr[i]],p2)

  y_ech2=ech2[,2]

  if(nh[gr[i]]!=0) {ty[i]=sum(y_ech2)*Ni[gr[i]]/nh[gr[i]]*(M/m)}

ss=c(0,0)

    Ni[gr[i]]

 for(q in 1:Ni[gr[i]])

   {

   for(p in 1:Ni[gr[i]])

     {

     if(p!=q)

{ss[i]=ss[i]+(grappe[q]*grappe[p]*(M^2)*(Ni[gr[i]]^2)*(1-
1/(Ni[gr[i]]*M)))+((grappe[q]/nh[gr[i]]/Ni[gr[i]])^2*(1- 1/(M*Ni[gr[i]])))}

     }}}

  ybar_chap[l]=sum(ty)/100

  Ni_ech=as.vector(Ni[gr])

  ni_ech=round(Ni_ech/5)
```

```
 v1=sum(((ty*3)^2)*2/3)+2*((ty[1]/m/M)*(ty[2]/m/M)*(1- 1/(M*m)))

  v2=sum(ss)

 var_estim_ybar_Grappe=(v1+v2)/10000

 }

icSgmin=-1.96*sqrt(var_estim_ybar_Grappe)+ybar_chap


icSgmax=+1.96*sqrt(var_estim_ybar_Grappe)+ybar_chap

delta=icSgmax-icSgmin

SG2=data.frame(ybar_chap,var_estim_ybar_Grappe,icSgmin,icSgmax,delta)

s=1:10

h=SG2$ybar_chap

h1=SG2$icSgmin

h2=SG2$icSgmax

library(lattice)

xyplot(h2+h1+h ~ s,col=c("blue","red","green"),main="Sondage par grappes de second
degré",ylab="Age",xlab="",type='o')

mean(SG2$ybar_chap)

 grappetab=cbind(ybar_chap , var_estim_ybar_Grappe, icSgmin, icSgmax , delta )

colnames(grappetab)=c("ybar_chap", "var_estim_ybar_Grappe", "icSgmin", "icSgmax" , "delta")

View(grappetab)
```

les intervalles de confiances

```
library(gridExtra)

library(ggplot2)

library(lattice)

op <- par(mfrow=c(3,3),no.readonly = TRUE,new=T )


PESR1=plot(icPESRmin,ylim=c(min(min(icPESRmax),min(icPESRmin)),max(max(icPESRmax),max(i
```

```
cPESRmin))),type="o",xaxt ="n",xlab="les echantillons",ylab="Intervalle de confiance",col="blue",main
= "PESR")

par(new=T)

PESR=plot(icPESRmax,ylim=c(min(min(icPESRmax),min(icPESRmin)),max(

max(icPESRmax),max(icPESRmin))),type="o",col="deeppink",xaxt = "n",xlab="",ylab="")

PESR_tab=cbind(ybarPESR,varPESR,icPESRmin,icPESRmax,longueurIC,moy_PESR,
main="PESR")

a=plot(icPEARmin,ylim=c(min(min(icPEARmax),min(icPEARmin)),max(max(icPEARmax),max(icPEA
Rmin))),type="o",xaxt

    = "n",xlab="les ?chantillons",ylab="Intervalle de confiance",col="blue")

par(new=T)

PEAR=plot(icPEARmax,ylim=c(min(min(icPEARmax),min(icPEARmin)),max(max(icPEARmax),max(i
cPEARmin))),type="o",xaxt

= "n",xlab="les echantillons",ylab="Intervalle de confiance",cex=1.5,font=3,col="red",main="PEAR")

S=1:20

PIAR_IC=xyplot(q1+q2+q ~ s,col=c("blue","red","green"),main="PIAR",ylab="Age",type='o')

S=1:20

PISR=xyplot(l1+l2+l ~ s,col=c("blue","red","green"),main="PISR",ylab="Age",type='o')

p=plot(seq(1,20, by =1),ICmin,ylim=c(min(min(as.numeric(ICmax)),min(as.numeric(ICmin)))
```

```
, max(max(as.numeric(ICmax)),max(as.numeric(ICmin)))), type="o", xaxp=c(0,20,20) , col="red"
,xlab="Echantillons",ylab="Intervalles de confiance" , cex=1 , font=3)

par(new=TRUE)

strat_mm_taille=plot(seq(1,20, by =1),ICmax,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)),
max(as.numeric(ICmin)))),type="o",xaxp=c(0,20,20),col="blue",main = "strat_mm_taille")

p1=plot(seq(1,20, by =1),ICmin,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),
max(max(as.numeric(ICmax)),
max(as.numeric(ICmin)))),type="o",xaxp=c(0,20,20),
col="red",xlab="Echantillons",ylab="Intervalles de confiance",cex=1,font=3)

par(new=TRUE)

strat_all_prop=plot(seq(1,20, by =1),ICmax,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)),
max(as.numeric(ICmin)))),type="o",xaxp=c(0,20,20),col="blue",main ="strat_all_prop" )

p2=plot(seq(1,20, by =1),ICmin,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)), max(as.numeric(ICmin)))),
type="o",xaxp=c(0,20,20), col="red",xlab="Echantillons",ylab="Intervalles de confiance",cex=1,font=3)

par(new=TRUE)
```

```
strat_opt=plot(seq(1,20, by =1),ICmax,ylim=c(min(min(as.numeric(ICmax)),
min(as.numeric(ICmin))),max(max(as.numeric(ICmax)), max(as.numeric(ICmin)))),
type="o",xaxp=c(0,20,20), col="blue",xlab="Echantillons",ylab="Intervalles de
confiance",cex=1,font=3,main="strat_opt")


grappe_1=xyplot(h2+h1+h ~ s,col=c("blue","red","green"),main="Sondage par grappes de premier
degré",ylab="Age",xlab="",type='o')


grappe_2= xyplot(h22+h12+h222 ~ s,col=c("blue","red","green"),main="Sondage par grappes de
second degré",ylab="Age",xlab="",type='o')


par(op)
```