



Colloque Sondages

La pondération de l'enquête Sans Domicile 2012

Ensayi Novembre 2012

Lionel Viglino et Sylvain Quenum

Le vif du sujet :

Proba d'inclusion des contacts

$$\pi_a * m_k * \frac{F_s}{\sum_{s \in k} n_j * F_s} * \frac{nc}{\hat{F}_s} = \alpha * f * \frac{nc}{8} * \frac{F_s}{\hat{F}_s}$$

Le vif du sujet :

Poids des (questions des) QIL :
on répartit le poids des contacts francophones sur les
QIL

$$\frac{1}{\pi_a * m_k * \sum_{s \in k} \frac{F_s}{n_j * F_s} * \frac{nc}{\hat{F}_s}} * \frac{(1 - txNF) * nc}{nb \text{ QIL}}$$

Le vif du sujet :

Proba d'inclusion des (questions des) QIL :

$$\pi_a * m_k * \frac{F_s}{\sum_{s \in k} nj * F_s} * \frac{nb \text{ QIL}}{\hat{F}_s * (1 - txNF)}$$

Le vif du sujet :

Le nombre de Sans Domicile francophones :
avec les QIL car le semainier est de meilleure qualité

$$\sum_{(k,s,QIL)} \frac{\sum_{s \in k} nj * F_s}{\pi_a * m_k} * \frac{\hat{F}_s}{F_s} * \frac{(1 - txNF)}{nbQIL} * \frac{1}{nbliens}$$

Quelle base de sondage?

- ◆ par définition la population cible n'a pas de résidence propre fixe qu'on puisse sélectionner directement
- ◆ Il existe une solution théorique, le sondage indirect :
 - on échantillonne les lieux que fréquentent régulièrement les personnes ciblées pour arriver ensuite à y sélectionner les individus à enquêter
 - On applique le partage des poids (nbre de liens)



Exhaustivité?

- ◆ Recensement de tous les services d'aides gratuit du champ => enquête téléphonique
- ◆ Ambition d'exhaustivité des services, mais pas indispensable car on cible les usagers
- ◆ Taux de couverture assuré (pas de biais) si pour chaque individu du champ, il existe au moins un service qu'il fréquente dans notre base de services
- ◆ Collecte l'hiver

Coût limité?

- ◆ Coûts déplacement enquêteur, recensement services
- ◆ Tirage d'unités primaires :
 - 2 strates : tirage 80 agglomérations de plus de 20000 hbts et 80 agglomérations comprises entre 5000 et 20000 hbts
 - proportionnellement à leur population municipale 2006
 - seuil d'exhaustivité à 200 000 hbts
 - équilibrage sur la capacité Finess, le nombre de sans-abri estimé dans le recensement et les indicatrices des zones 'Sud', 'Est' et 'Reste'

Cœur de cible des individus

- ◆ Les francophones:
 - Questionnaire long QIL face à face en priorité
 - mais inclusion des non francophones dans le champ de l'enquête réédition 2012
 - Questionnaire auto-administré QAA court traduit
- ◆ Les sans-abri:
 - surreprésentation (5/3) des services de repas extérieurs, haltes de nuit et plans grand froid

Les non francophones

- ◆ modes de sélection, de questionnement et donc de pondération différenciés pour les francophones F et non francophones NF
 - Dès lors, besoin de connaître les nombres - ou du moins, les proportions - de francophones au sein des services d'aide
- ◆ Proportion NF non connue à l'avance:
 - au moment du tirage des services x jours
 - besoin d'estimation terrain lors de la collecte

Les services x jours

- ◆ 80 agglomérations > 20000 hbts:
 - 2 vides de service (dont monaco!), 3 ont trop peu
 - 75 agglos conservées, repondérées
- ◆ 80 agglomérations entre 5000 et 20000 hbts:
 - 29 avec au moins 1 service
 - Collecte Ined que dans 13 agglos, inférence à voir
- ◆ Unités secondaires :
 - Tirage de 1300 services x jours, stratifié
 - Proportionnellement à la fréquentation F+NF
 - Tirage d'une visite par service pour les < 20000

Tirage des individus

- ◆ Le troisième degré : tirage en 2 phases
 - sélection d'un certain nombre d'individus - dits contacts - selon la taille du service (en usagers) et dans la limite de 16. Ils servent à estimer la proportion de non francophones et à fournir des données de calage (sexe)
 - puis, sélection de personnes francophones parmi les individus de la première phase, dans la limite de 4 => QIL, QAA pour autres

Exemple table de tirage

Effectif entre 173 et 228

| | | | | | |
|---------------------------------|-------------------|---------------------|------------------------|------------------------|-----|
| Rangs de contacts obligatoires | 3 17 32 45 | 60 74 88 102 | 116 131 145 159 | 173 187 202 215 | +58 |
| Rangs éventuels de remplacement | | | | | |

Effectif entre 229 et 303

| | | | | | |
|---------------------------------|-------------------|----------------------|------------------------|------------------------|-----|
| Rangs de contacts obligatoires | 4 23 42 60 | 79 98 117 135 | 154 173 192 210 | 229 248 267 285 | +75 |
| Rangs éventuels de remplacement | | | | | |

Effectif entre 304 et 401

| | | | | | |
|---------------------------------|-------------------|------------------------|------------------------|------------------------|-----|
| Rangs de contacts obligatoires | 5 30 56 79 | 105 130 155 179 | 204 229 254 278 | 303 328 353 377 | +99 |
| Rangs éventuels de remplacement | | | | | |

Pondération théorique

- Probabilité inclusion au niveau contact(3 degrés)

$$\pi^T_{takv_s j} = \pi_{ta} \times \pi_{v_s/ta} \times \pi_{j/v_s ta} = \pi_{ta} \times \left[m_{tak} \times \frac{Freq_{taks}}{\sum_{r \in k} (nj_{takr} \times Freq_{takr})} \right] \times \frac{nc_{taks}}{Freq_{taks}}$$

- Poids contact (1ere phase):

$$P^T_{takv_s j} = \frac{\sum_{r \in k} (nj_{takr} \times Freq_{takr})}{\pi_{ta} \times m_{tak} \times nc_{taks}}$$

- Poids QIL (1ere et 2eme phases)

$$P^T_{takv_s ji} = \left[\frac{\sum_{r \in k} (nj_{takr} \times Freq_{takr})}{\pi_{ta} \times m_{tak} \times nc_{taks}} \right] \times \left[\frac{(1 - p_{taks}) \times nc_{taks}}{nqil_{taks}} \right]$$

Recherche équi pondération

- ◆ Au sein des agglos de plus de 20000 hbts
- ◆ Lors tirage des services x jours (m inconnu)
 - Niveau contact ou francophone?
 - Cœur de cible francophones, mais p non encore connu
 - Donc mise en œuvre niveau contact
 - $P_0 = \text{cste}$ (au facteur alpha de surreprésentation près)

$$P^T_{takv_s j} = \frac{P_0}{\alpha_{tak}} \frac{\sum_{r \in k} (nj_{takr} \times Freq_{takr})}{\pi_{ta} \times m_{tak} \times nc_{taks}} = \frac{P_0}{\alpha_{tak}}$$

Calcul des allocations m

- On peut fixer un nombre moyen théorique de 8 contacts (les contacts réels variant entre 1 et 16), d'où:

$$m_{tak} = \frac{\alpha_{tak} \times \sum_{r \in k} (nj_{takr} \times Freq_{takr})}{P_0 \times \pi_{ta} \times 8}$$

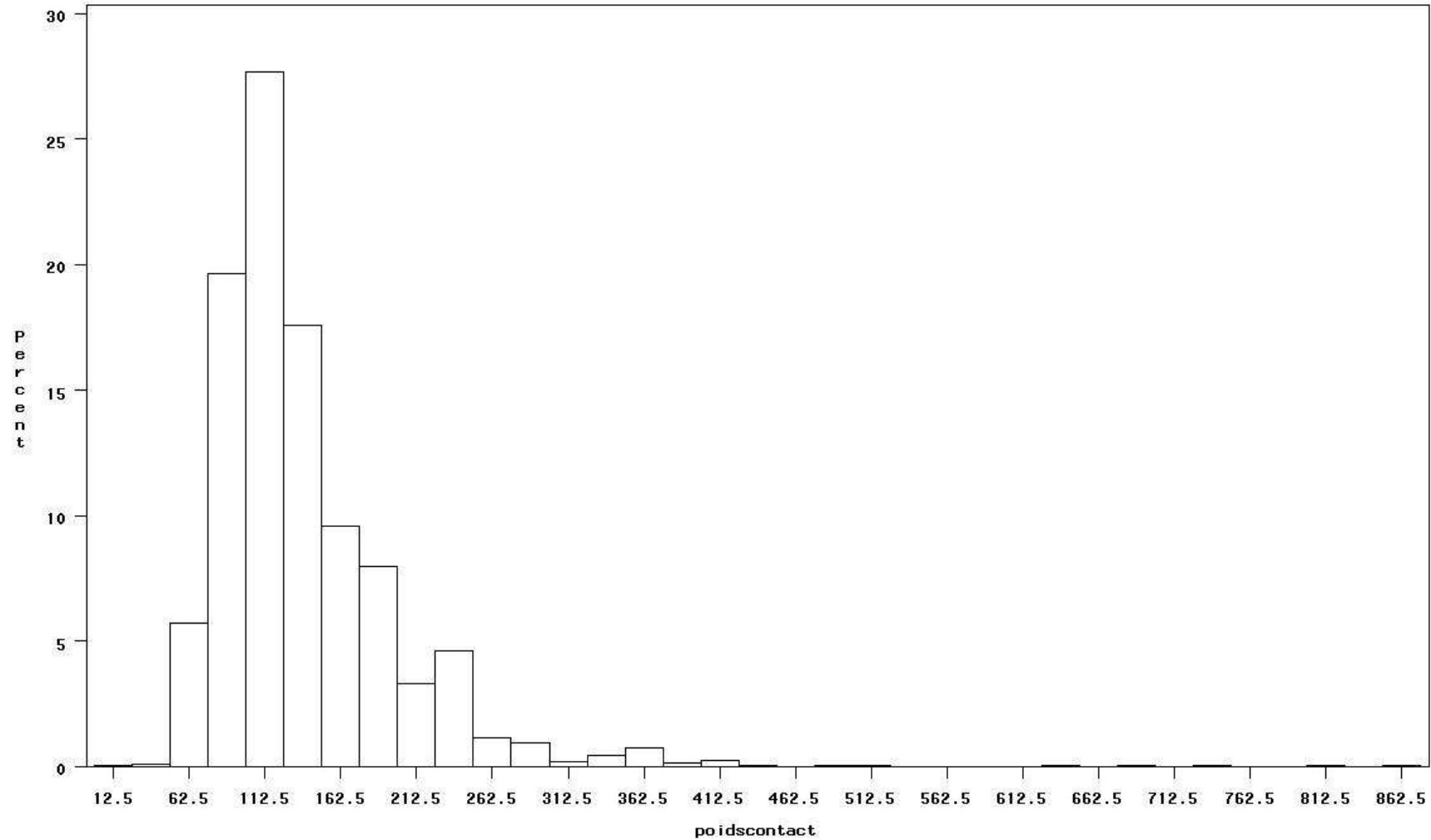
$$\sum_{k \in ta} m_{tak} = 1300$$

$$\sum_{k \in ta} \left[\frac{\alpha_{tak} \times \sum_{r \in k} (nj_{takr} \times Freq_{takr})}{P_0 \times \pi_{ta} \times 8} \right] = \frac{\sum_{k \in ta} \left[\frac{\alpha_{tak} \times \sum_{r \in k} (nj_{takr} \times Freq_{takr})}{\pi_{ta}} \right]}{P_0 \times 8} = 1300$$

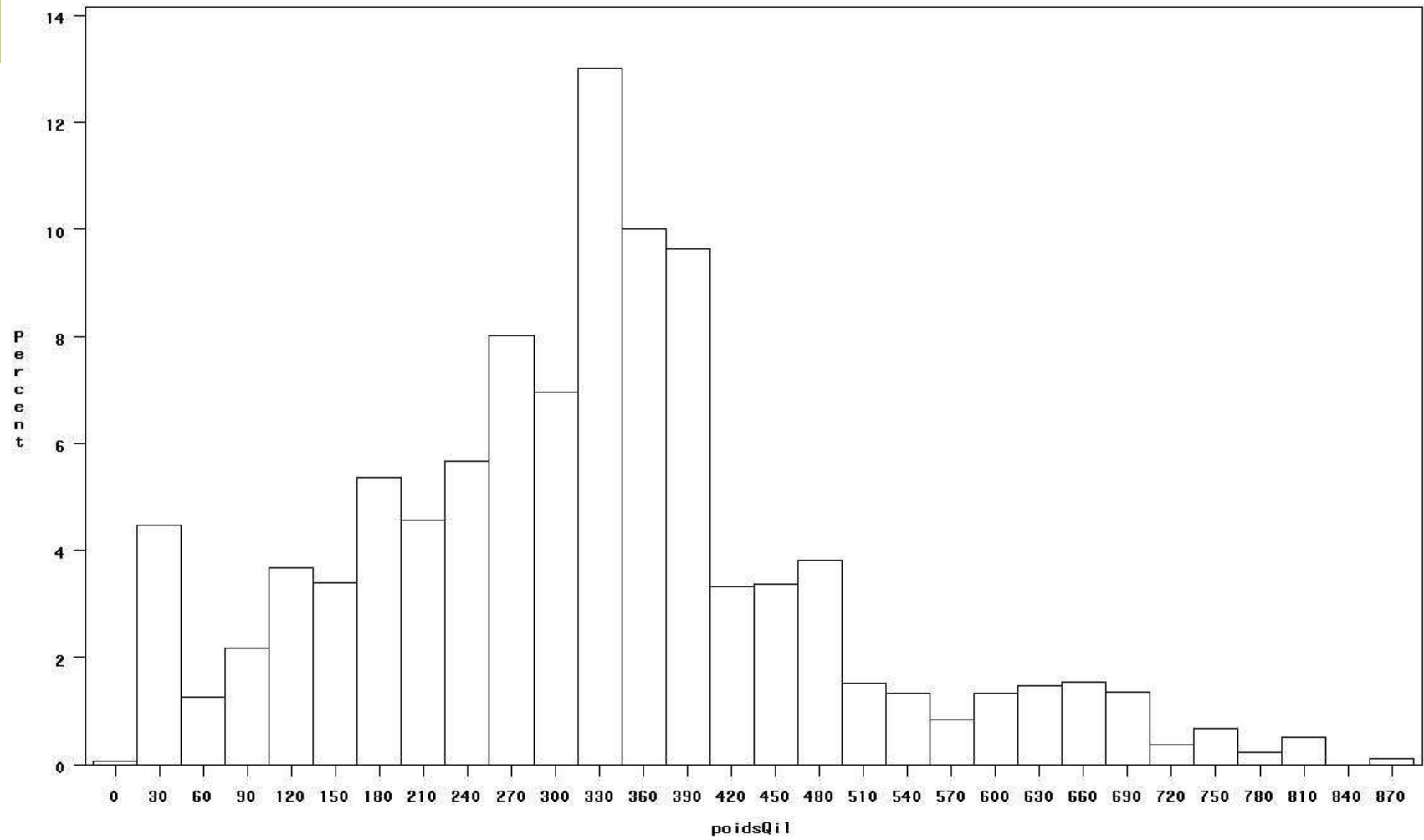
$$P_0 = \frac{\sum_{k \in ta} \left[\frac{\alpha_{tak} \times \sum_{r \in k} (nj_{takr} \times Freq_{takr})}{\pi_{ta}} \right]}{1300 \times 8}$$

$$P_0 \approx 200$$

poids contacts



poids QIL



| TYPE_SERVICE | nb prestations | | | |
|---|----------------|---------------|----------------|---------------|
| | univers | % | ech | % |
| Accueil_nuit | 25289 | 1,1% | 20776 | 0,9% |
| Autre_hébergement_dispersé | 210315 | 8,9% | 197295 | 8,3% |
| Autre_hébergement_en_hôtel_FJT_FTM_RS | 57805 | 2,4% | 50044 | 2,1% |
| Autre_hébergement_regroupé | 194246 | 8,2% | 197780 | 8,4% |
| Centre_maternel_en_hébergement_dispersé | 24193 | 1,0% | 19508 | 0,8% |
| Centre_maternel_en_hébergement_regroupé | 35213 | 1,5% | 27537 | 1,2% |
| Centre_maternel_en_hôtel_FJT_FTM_RS | 1886 | 0,1% | 1582 | 0,1% |
| Hébergement_dispersé_insertion | 249404 | 10,5% | 252419 | 10,7% |
| Hébergement_dispersé_stabilisation | 28049 | 1,2% | 24833 | 1,0% |
| Hébergement_dispersé_urgence | 82202 | 3,5% | 89397 | 3,8% |
| Hébergement_en_hôtel_FJT_FTM_RS_insertion | 9373 | 0,4% | 6769 | 0,3% |
| Hébergement_en_hôtel_FJT_FTM_RS_stabilisation | 4328 | 0,2% | 1560 | 0,1% |
| Hébergement_en_hôtel_FJT_FTM_RS_urgence | 41336 | 1,7% | 43392 | 1,8% |
| Hébergement_regroupé_insertion | 233989 | 9,9% | 232703 | 9,8% |
| Hébergement_regroupé_stabilisation | 102549 | 4,3% | 110922 | 4,7% |
| Hébergement_regroupé_urgence | 220775 | 9,3% | 234272 | 9,9% |
| Restauration_midi_extérieur | 3272 | 0,1% | 2896 | 0,1% |
| Restauration_midi_intérieur | 262090 | 11,1% | 257050 | 10,9% |
| Restauration_ptdej_extérieur | 3402 | 0,1% | 3783 | 0,2% |
| Restauration_ptdej_intérieur | 248402 | 10,5% | 256640 | 10,8% |
| Restauration_soir_extérieur | 104419 | 4,4% | 107737 | 4,6% |
| Restauration_soir_intérieur | 126598 | 5,4% | 130237 | 5,5% |
| SAMU Hébergement_en_hôtel_FJT_FTM_RS_urgence | 97040 | 4,1% | 97040 | 4,1% |
| Total | 2366174 | 100,0% | 2366174 | 100,0% |

| TYPE_SERVICE | nb francophones | | | |
|---|-----------------|---------------|----------------|---------------|
| | univers | % | ech | % |
| Accueil_nuit | 18221 | 1,0% | 15215 | 0,9% |
| Autre_hébergement_dispersé | 163758 | 9,3% | 154022 | 8,7% |
| Autre_hébergement_en_hôtel_FJT_FTM_RS | 40205 | 2,3% | 35701 | 2,0% |
| Autre_hébergement_regroupé | 157803 | 8,9% | 157928 | 9,0% |
| Centre_maternel_en_hébergement_dispersé | 22677 | 1,3% | 17181 | 1,0% |
| Centre_maternel_en_hébergement_regroupé | 31429 | 1,8% | 23943 | 1,4% |
| Centre_maternel_en_hôtel_FJT_FTM_RS | 1348 | 0,1% | 1266 | 0,1% |
| Hébergement_dispersé_insertion | 213221 | 12,1% | 217435 | 12,3% |
| Hébergement_dispersé_stabilisation | 24487 | 1,4% | 22147 | 1,3% |
| Hébergement_dispersé_urgence | 66477 | 3,8% | 73565 | 4,2% |
| Hébergement_en_hôtel_FJT_FTM_RS_insertion | 6543 | 0,4% | 3886 | 0,2% |
| Hébergement_en_hôtel_FJT_FTM_RS_stabilisation | 1932 | 0,1% | 156 | 0,0% |
| Hébergement_en_hôtel_FJT_FTM_RS_urgence | 19900 | 1,1% | 24039 | 1,4% |
| Hébergement_regroupé_insertion | 203177 | 11,5% | 205076 | 11,6% |
| Hébergement_regroupé_stabilisation | 80898 | 4,6% | 90207 | 5,1% |
| Hébergement_regroupé_urgence | 153408 | 8,7% | 157956 | 9,0% |
| Restauration_midi_extérieur | 2612 | 0,1% | 2102 | 0,1% |
| Restauration_midi_intérieur | 171565 | 9,7% | 165737 | 9,4% |
| Restauration_ptdej_extérieur | 561 | 0,0% | 378 | 0,0% |
| Restauration_ptdej_intérieur | 168573 | 9,5% | 170858 | 9,7% |
| Restauration_soir_extérieur | 66408 | 3,8% | 70788 | 4,0% |
| Restauration_soir_intérieur | 87190 | 4,9% | 87023 | 4,9% |
| SAMU Hébergement_en_hôtel_FJT_FTM_RS_urgence | 65016,8 | 3,7% | 65017 | 3,7% |
| Total | 1767411 | 100,0% | 1761625 | 100,0% |

Pondération réelle

- Probabilité inclusion au niveau contact(3 degrés)

$$\pi^{R}_{takv_s j} = \pi_{ta} \times m_{tak} \times \frac{Freq_{taks}}{\sum_{r \in k} (nj_{takr} \times Freq_{takr})} \times \frac{nc_{takv_s}}{Freq^R_{takv_s}}$$

- Poids contact (1ere phase):
- $$P^{R}_{takv_s j} = \frac{Freq^R_{takv_s} \times \sum_{r \in k} (nj_{takr} \times Freq_{takr})}{\pi_{ta} \times m_{tak} \times nc_{takv_s} \times Freq_{taks}}$$

- Poids QIL (1ere et 2eme phases)

$$P^{R}_{takv_s ji} = \left[\frac{Freq^R_{takv_s} \times \sum_{r \in k} (nj_{takr} \times Freq_{takr})}{\pi_{ta} \times m_{tak} \times nc_{takv_s} \times Freq_{taks}} \right] \times \left[\frac{(1 - p^{R}_{takv_s}) \times nc_{takv_s}}{nqil_{takv_s}} \right]$$

Redressements

- Services : défaut et refus : règle de trois
- QIL + QAA avec la feuille de contact pour :
H/F et F/NF car les taux de refus sont peut être différents
- QIL avec QIL + QAA pour les variables communes.
- Individu : non car liens encore inconnus



Partage des poids

- ◆ Un individu peut fréquenter plusieurs services d'aides différents voire même être enquêté plusieurs fois
- ◆ On diffusera le nombre d'utilisateurs et non pas de prestations servies
- ◆ On collecte le semainier pour calculer le nombre de fois qu'un individu a fréquenté un service de la base de sondage sur les 4 semaines (nombre de liens)

réinterrogations

- ◆ Pour un francophone tiré pour un QIL plusieurs fois, on lui passe un QIC les fois suivantes:
 - On recollecte la partie qui varie dans le temps (dont semainier) => 2 systèmes de poids selon les variables
 - On a des informations : date naissance, sexe, ... pour apparier les questionnaires du même individu
 - Pas faisable pour les QAA



Calcul poids final individus

- ◆ $\text{Poids} = \frac{\text{somme (poids interrogations)}}{\text{nb liens dans la base de sondage}}$

dispersion des poids

- ◆ Pendant la collecte, les poids des services jours sont modifiés du fait de fusions, éclatements, rééchantillonnages de visites.
- ◆ L'équipondération des contacts visée initialement est donc perturbée par :
 - Ces modifications
 - Des fréquentations bien plus faibles lors de la visite
 - Le partage des poids



◆ **MERCI DE VOTRE ATTENTION**