

L'utilisation des historiques d'appels pour redresser une enquête téléphonique: une étude par simulation à partir de l'enquête Fecond

Stéphane Legleye¹, Nirintsoa Razakamanana², Géraldine Charrance² et Hélène Juillard²

1. INED; 2. Inserm, U1178; Univ Paris-Sud; Univ Paris Descartes, UMR-S1178

2. INED

Journées de méthodologie statistique, 2015



Problématique – contexte (1)

- Téléphone théoriquement efficace pour les enquêtes en population générale
 - Complexité de l'offre téléphonique (box, portables, fixes, pas d'annuaires) :
 - numérotation aléatoire, probabilités d'inclusion multiples
 - Complexité de l'usage du téléphone : filtrage des numéros appelants suivant les équipements
 - Existence d'échanges de type « machine à machine »
- Baisse du taux de participation depuis quelques années
 - Lassitude, confusion des enquêtes de recherche avec les études marketing, sollicitation intense des abonnés

Problématique – contexte (2)



- Les redressements font toujours appel à un calage sur marges
- Des parodonnées disponibles pour tous les numéros appelés:
(Historiques des appels: heures, jours et issues des appels)
 - Bien mesurées (ordinateur, automate d'appel CATI)
 - Elles sont fortement liées à la participation à l'enquête, par construction
 - Elles présentent des liens faibles mais mesurables avec les variables d'intérêt
 - Il est inhabituel de les utiliser:
 - Données de gestion du terrain téléphonique
 - Nécessité de traitements préalables importants

- CNRT utile si :
 - Lien entre variable auxiliaire et participation fort
 - Mais peut augmenter la variance si pas de lien avec variables d'intérêt
 - Lien entre variable auxiliaire et variables d'intérêt est fort
 - Mais peut augmenter la variance si pas de lien avec la participation

- Tester l'usage de paradonnées pour une correction de la non-réponse totale (CNRT)
 - Simulation de plusieurs mécanismes de non-réponse
 - Utilisation de plusieurs jeux de paradonnées
 - Comparaison de plusieurs méthodes de CNRT avant calage et comparaison avec le calage direct et l'absence de redressement

Cadre de travail



- L'échantillon des répondants de l'enquête est la population cible
- Pas de sondage, mais non-réponse totale (NRT), générée suivant plusieurs mécanismes
- CNRT par groupes de réponses homogènes (GRH)
- Simulations: écart quadratique moyen empirique (MSE) calculé pour toutes les méthodes, tous les mécanismes de NRT et 9 variables d'intérêt

L'enquête



- Enquête Fecond (Fécondité et dysfonctions sexuelles), Inserm-Ined, 2010
- Taux de réponse 44,8%, n=8639
- Champ : 15-49 ans francophones métropole, ménages ordinaires

Les données



- 8 variables sociodémographiques **SD** :
 - Sexe → lieu de naissance (en France, à l'étranger)
 - âge (7 tranches d'âge) → situation professionnelle (emploi ou non)
 - diplôme le plus élevé (5 niveaux) → vie de couple (en couple ou non)
 - zone géographique de résidence (4 catégories) → taille du ménage (1 pers., 2, 3-4, 5 et plus)
- 9 variables d'intérêt binaires **VI**:
 - rapport sexuel avec un partenaire du sexe opposé ; du même sexe
 - rapport sexuel douze derniers mois
 - cinq partenaires sexuels ou plus du sexe opposé ; du même sexe
 - dernier rapport sexuel avec le partenaire régulier
 - même partenaire pour toutes les grossesses
 - relation sexuelle imposée au cours des douze derniers mois
 - avoir causé ou avoir vécu une interruption volontaire de grossesse

Les parodonnées



4 issues d'appel :

- non-contact
- refus
- rendez-vous
- passation (partielle ou complète)

Jours et heures d'appel :

→ 3 plages horaires par jour, 3 types de jours de semaine → 9 modalités

- **P1**: bilan anhistorique des issues : 4 variables
- **P2**: bilan ventilé par plages horaires: 36 variables
- **P3**: trajectoire d'appel: succession des issues sur 50 appels (maximum), en ajoutant une issue fictive (non-concerné)

Liens entre paradonnées et données (population totale: n=8639) (1)



▲ ACP sur P1 et P2, ACP sur SD et VI, analyse harmonique qualitative (AHQ) sur P3. Examen des 3 premiers axes.

- Corrélations fortes entre les paradonnées
 - Premiers axes en particulier ($\rho > 0,66$)
- Corrélations modérées entre P_i et SD (axes 2 et 3 des P_i , axe 2 des SD : $\rho \approx \# 0,15$)
- Corrélations faibles entre P_i et VI ($\rho < 0,06$)
- Corrélations modérées à fortes entre SD et VI ($\rho \in [0,04; 0,35]$)

Liens entre paradonnées et données (population totale: n=8639) (2)



- Une analyse de régression logistique des 9 VI montre que les répondants difficiles (plus de 20 appels ou un refus) ont
 - des caractéristiques SD et des VI particulières
 - la différence sur les VI n'est pas réductible à l'ajustement sur les SD
- Des régressions linéaires des axes principaux des VI sur les axes principaux des Pi et des SD montre que
 - les coefficients associés aux Pi sont significatifs quoique faibles (b max=0,1 pour les axes 1 et 3 de P3)
 - Mais les variances expliquées par les Pi sont faibles (environ 3%)

Cadre de travail

Cadre de la simulation sur les répondants

- défavorable
 - lien P_i – VI faible
- mais potentiellement efficace
 - Lien P_i – non réductible aux SD

En vrai, les NRT sont particuliers du point de vue des VI, des SD et des P_i : c'est ce qui va être exploité

5 Mécanismes de non-réponse

- MAR1 : quintiles sur l'axe principal 1 des SD
- MAR2 : quintiles sur l'axe principal 1 de P1
- MAR3 : quintiles sur la somme des deux précédents
- NMAR : quintiles sur l'axe principal 1 des VI

taux de NRT de 0,35 à 0,75 → taux de NRT # 50%

- MCAR : totalement au hasard
- 1000 réplifications **pour chaque mécanisme**

Constitution des GRH

- Soit régression logistique de la participation (P1 et P2)
- Soit ACP (P1 et P2) ou AHQ (P3) puis classification (K-means)

Sur tous les mécanismes et tous les P_i

Nombre de GRH=7

Méthodes comparées

- 1: acP_P1 : ACP et classification sur P1 (bilan anhistorique)
- 2: log_P1 : régression logistique sur P1
- 3: acP_P2 : ACP sur P2 (plages horaires)
- 4: log_P2 : régression logistique P2 (plages horaires)
- 5: ahq_P3 : AHQ sur P3 (trajectoires d'appel)

- Puis calage sur les 8 variables SD → 10 méthodes

2 méthodes témoins :

- Calage direct
- Aucun traitement

Calcul des performances

-
- $MSE_{Emp}(\hat{y}) = (BR_{Emp}(\hat{y}))^2 + Var_{Emp}(\hat{y})$
- $MSE_{Emp}(\hat{y}) = \left[\frac{1}{1000} \sum_{b=1}^{1000} \frac{\hat{y}^b - \bar{y}}{\bar{y}} \right]^2 + \frac{1}{1000} \sum_{b=1}^{1000} (\hat{y}^b - \hat{y})^2.$

Erreur quadratique moyenne (MSE)



	Meilleur (2 ^{ème})	Gain relatif/ aucun traitement	Gain relatif/ calage direct
MAR1 : SD	Calage_log_P2	30.1%	3.8%
	(Calage_log_P1)	(29.0%)	(2.3%)
MAR2 : P1	Log_P1	52.7%	45.9%
	(Calage_log_P1)	(52.4%)	(45.6%)
MAR3 : SD – P1	Calage_log_P1	43.2%	16.1%
	(Calage_log_P2)	(41.1%)	(13.0%)
MCAR	Calage_acp_P2	5.1%	0.0%
NMAR : VI	Calage_log_P2	10.3%	1.1%
Moyenne sur l'ensemble des mécanismes	Calage_log_P1	20.0%	5.3%
	(Calage_log_P2)	(19.9%)	(5.1%)

Bilan (1)

- Le calage direct n'est jamais la meilleure option
 - Même lorsque le mécanisme de NRT est fondé sur les variables du calage (MAR1=3,8%)
 - Même lorsque le mécanisme est NMAR
 - Le gain est fort lorsque le mécanisme est fondé sur les parodonnées (MAR2) ou un mixte de parodonnées et de SD (MAR3)

Bilan (2)

- La méthode du score est toujours supérieure à la classification (ACP ou AHQ) (sauf MCAR)
 - Déjà vu dans des études de simulation
- Le jeu de parodonnées utilisé pour les GRH a une faible importance
 - Dû à la communauté d'information entre les Pi

Bilan (3)

- Résultats favorables d'autant plus que le lien entre P_i et participation est très faible dans nos simulations (sauf MAR2 et MAR3)
 - Pas d'augmentation de variance
- Une macro SAS paramétrable est disponible
 - Les mécanismes et taux de NRT sont paramétrables
- Une version contrôlant mieux les liens NRT – P_i et NRT – VI est en cours
 - Outil d'aide à la décision pour une CNRT ad hoc.

Références

- Beck, F., S. Legleye, et al. (2005). "Aux abonnés absents : liste rouge et téléphone portable dans les enquêtes en population générale sur les drogues." Bulletin de méthodologie sociologique **n°86**(Avril 2005): 5-29.
- Haziza, D. and J.-F. Beaumont (2007). "On the Construction of Imputation Classes in Surveys." International statistical review **75**(1): 25-43.
- Legleye, S., G. Charrance, et al. (2013). "Improving survey participation: cost effectiveness of call-backs to refusals and increased call attempts in a national telephone survey in France." Public Opinion Quarterly **77**(3): 666-695.
- Little, R. J. A. and S. Vartivarian (2005). "Does weighting for nonresponse increase the variance of survey means?" Survey methodology **31**(2): 161–168.
- Maitland, A., C. C. Cordero, et al. (2009). An exploration into the use of paradata for nonresponse adjustment in a health survey. JSM proceedings. Alexandria, VA, American Statistical Association: 370-378.