



*Cours et cas pratiques*

**Master • Écoles d'ingénieurs**

# MÉTHODES DE SONDAGE

- ▶ **Enquêtes électorales**
- ▶ **Enquêtes dans le domaine de la santé**
- ▶ **Enquêtes dans les pays en développement**

*Sous la direction de*  
**Philippe Guilbert, David Haziza,  
Anne Ruiz-Gazen et Yves Tillé**

# MÉTHODES DE SONDAGE

- ▶ Enquêtes électorales
- ▶ Enquêtes dans le domaine de la santé
- ▶ Enquêtes dans les pays en développement

# Consultez nos parutions sur dunod.com

Dunod Éditeur, édition de livres, Microsoft Press, ETSF, Ediscience, InterEditions

http://www.dunod.com/

Dunod Éditeur, édition de livres, Microsoft Press, ETSF, Ediscience, InterEditions

Recherche  OK Collections Index thématique

Mon compte


 Ediscience  
 ETSF  
 InterEditions  
 Microsoft Press

LES RUBriques

Accueil Contacts

Interviews

Réinventer les RH : urgence !  
Giles Verrier

Ramassez 2008 : exigez la nouvelle formule !  
Thierry de Montbrial

→ toutes les interviews

Club Enesdynamis

→ Inscrivez-vous !

Événements

Découvrez le **videoblog**  
*Profession dirigeant*

En librairie ce mois-ci

Développement personnel et coaching : découvrez le NOUVEAU SITE [intereditions.com](http://intereditions.com) !

les librairies

Sciences et Techniques

Informatique

Gestion et Management

Sciences Humaines

Acheter  Mon panier

LES BIBLIOTHEQUES DES METIERS

→ Bibliothèque du DSI

→ Gestion industrielle

→ Métiers de la vigne et du vin

→ Marketing et Communication

→ Directeur d'établissement social et médico-social

→ Toutes les bibliothèques

LES NEWSLETTERS

→ Action sociale

→ Psychologie

→ Développement personnel et Bien-être

→ Entreprise

→ Expertise comptable

→ Informatique et NTIC

→ Industrie

→ Toutes les newsletters

Bacchus 2008  
 Dr. B.K. Bhatnagar et  
 principes de la  
 chimie  
 Jean Pierre  
 Couderc, Hervé  
 Harnin,  
 François  
 d'Hautville,  
 Etienne  
 Montagne

Python

Python  
 Petit guide de la  
 programmation  
 Tarek Zade

150 questions-answers  
 à la psychologie  
 et sport

150 questions-answers  
 à la psychologie et sport  
 Pascal F. Fournier  
 et  
 Yves Fournier

Profession dirigeant  
 De la  
 conception du  
 changement à  
 l'action  
 Gérard Roth,  
 Michel Kuryla

bibliothèques des métiers    newsletters    Microsoft Press    ediscience.net    expert-sup.com  
 Notice légale

# MÉTHODES DE SONDAGE

- Enquêtes électorales
- Enquêtes dans le domaine de la santé
- Enquêtes dans les pays en développement

## Cours et cas pratiques

Sous la direction de

***Philippe Guilbert***

Directeur adjoint des affaires scientifiques à  
L'Institut national de prévention et d'éducation pour la santé

***David Haziza***

Professeur adjoint à l'Université de Montréal

***Anne Ruiz-Gazen***

Maître de conférences à l'Université Toulouse 1

***Yves Tillé***

Professeur à l'Université de Neuchâtel

DUNOD

## Illustration de couverture : *Digitalvision*<sup>®</sup>

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1<sup>er</sup> juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée. Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, Paris, 2008  
ISBN 978-2-10-053829-4

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

# Table des matières

## AVANT-PROPOS

## CHAPITRE 1 • QUALITÉ DES ENQUÊTES ET STATISTIQUE PUBLIQUE

|    |                                                                                                                                                                                    |    |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1. | Remarques sur l'estimation à partir de sources multiples<br><i>Jean-Michel Charpin</i> .....                                                                                       | 1  |
| 2. | Défis liés à la mise en place d'outils de gestion de la qualité dans les enquêtes<br><i>France Lapointe, Jimmy Baulne et Louise Bourque</i> .....                                  | 6  |
| 3. | Eurobaromètres, EVS, ISSP, ESS : quatre enquêtes internationales<br><i>Pierre Bréchon</i> .....                                                                                    | 11 |
| 4. | Développements en cours du projet OCTOPUSSE de nouvel<br>échantillon-maître de l'Insee<br><i>Marc Christine et Sébastien Faivre</i> .....                                          | 16 |
| 5. | Vérification sélective des données pour le futur système français de statistiques<br>structurelles d'entreprises<br><i>Gwennaëlle Brilhault et Philippe Brion</i> .....            | 21 |
| 6. | Garantir la qualité de données d'entreprises françaises pour une enquête européenne<br><i>Agnès Checcaglini, Isabelle Recotillet et Patrick Rousset</i> .....                      | 26 |
| 7. | Comparaison des déclarations des salariés et des employeurs dans<br>l'enquête Familles-Employeurs<br><i>Nicolas Razafindratsima, Anne Papadopoulos et Bernard De Clédats</i> ..... | 31 |

## CHAPITRE 2 • ENQUÊTES LONGITUDINALES

|    |                                                                                                                                                                                                      |    |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1. | Le traitement combiné des effets de non-réponse non-ignorable et de sondage informatif dans l'analyse des données issues des enquêtes longitudinales<br><i>Gad Nathan et Abdulhakeem Eideh</i> ..... | 37 |
| 2. | L'imputation dans les enquêtes longitudinales au Canada<br><i>Michelle Simard</i> .....                                                                                                              | 45 |
| 3. | Calcul de la précision des estimations longitudinales dans l'Enquête Emploi en Continu<br><i>Dominique Place</i> .....                                                                               | 53 |
| 4. | Mise en place d'un système d'observation longitudinale des enfants en danger<br><i>Lucy Marquet, Juliette Halifax, Emmanuelle Guyavarch, Olivier Cohen et Catherine Quantin</i> .....                | 58 |
| 5. | Le panel d'élèves entrés en 6 <sup>ème</sup> en 1995<br><i>Nathalie Caron et Sylvie Lemaire</i> .....                                                                                                | 63 |
| 6. | Étude de l'attrition du panel suisse de ménages : applications à l'enquête SILC en Suisse<br><i>Eric Graf</i> .....                                                                                  | 67 |
| 7. | Impact de l'attrition d'un panel de jeunes conducteurs sur l'analyse du risque<br><i>Sylvain Lassarre, Cécile Coquelet, Pierre-Alain Hoyau et Agénor Lahatte</i> .....                               | 72 |

## CHAPITRE 3 • ENQUÊTE ET SANTÉ

|     |                                                                                                                                                                                                                                                                                       |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1.  | Les cohortes épidémiologiques prospectives : principes, difficultés, solutions envisagées<br><i>Marcel Goldberg, Alice Gueguen, Annette Leclerc, Jean-Louis Lanoé et Marie Zins</i> .....                                                                                             | 79  |
| 2.  | Enquêter les soignés<br><i>Patrick Peretti-Watel</i> .....                                                                                                                                                                                                                            | 87  |
| 3.  | Enquête de prévalence du VIH et du VHC, complétée par une recherche socio-anthropologique chez les usagers de drogues en France métropolitaine, 2004-2007<br><i>Marie Jauffret-Roustide, Yann Le Strat, Nicolas Razafindratsima, Julien Emmanuelli et Jean-Claude Desenclos</i> ..... | 94  |
| 4.  | Les nouveaux développements dans les enquêtes sur la santé au Canada<br><i>Michelle Simard et Johane Dufour</i> .....                                                                                                                                                                 | 99  |
| 5.  | Événements de vie et santé : principes et méthodes<br><i>Catherine Cavalin</i> .....                                                                                                                                                                                                  | 108 |
| 6.  | Le recours à un répondant Proxy importe-t-il dans les enquêtes sur la santé ?<br><i>Béregère Davin, Xavier Joutard, Alain Paraponaris et Pierre Verger</i> .....                                                                                                                      | 113 |
| 7.  | Les évolutions méthodologiques d'une enquête répétée : avantages et inconvénients. L'exemple du Baromètre santé<br><i>Arnaud Gautier, Christophe Léon, François Beck et Philippe Guilbert</i> .....                                                                                   | 118 |
| 8.  | L'Enquête suisse sur la santé : présentation et méthodologie d'une grande enquête nationale de santé<br><i>Renaud Lieberherr</i> .....                                                                                                                                                | 122 |
| 9.  | Approche biographique de l'espace de vie, de la santé et des événements marquants auprès de seniors vivant en Belgique<br><i>Laurence Thomsin, Yannick Banturiki, Nicolas Rogister et Ingrid Schockaert</i> .....                                                                     | 127 |
| 10. | La mesure des consommations alimentaires par l'étude INCA2<br><i>Ariane Dufour, Lionel Lafay et Jean-Luc Volatier</i> .....                                                                                                                                                           | 132 |

**CHAPITRE 4 • SONDAGES ÉLECTORAUX**

|    |                                                                                                                                                                                  |     |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1. | La commission des sondages et les campagnes électorales<br><i>Elisabeth Dupoirier</i> .....                                                                                      | 139 |
| 2. | Délibération et changement d'opinion lors d'un sondage<br><i>Patrick Fournier, Mathieu Turgeon, André Blais, Elisabeth Gidengil,<br/>Neil Nevitte et Joanna Everitt</i> .....    | 147 |
| 3. | L'espace politique français à la veille de l'élection présidentielle (2007) :<br>analyse géométrique des données<br><i>Jean Chiche, Viviane Le Hay et Brigitte Le Roux</i> ..... | 152 |
| 4. | La détermination par les instituts de l'indétermination des électeurs<br><i>Patrick Lehingue</i> .....                                                                           | 158 |
| 5. | Les probabilités de vote<br><i>Bernard Denni et Philippe Caillot</i> .....                                                                                                       | 164 |
| 6. | L'appui à la souveraineté du Québec : effets de formulation de la question et<br>effets de contexte<br><i>François Yale et Claire Durand</i> .....                               | 169 |

**CHAPITRE 5 • SONDAGES DANS LES PAYS EN DÉVELOPPEMENT**

|     |                                                                                                                                                                                                                                                          |     |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1.  | L'estimation du nombre d'enfants victimes de traite au Bénin :<br>un défi méthodologique majeur<br><i>Blandine C. Dansou</i> .....                                                                                                                       | 175 |
| 2.  | La pratique des enquêtes par sondages auprès des ménages au Liban<br><i>Najwa Yaacoub</i> .....                                                                                                                                                          | 180 |
| 3.  | Gouvernance, démocratie et lutte contre la pauvreté : enseignements tirés des enquêtes<br>officielles auprès des ménages en Afrique francophone et dans la Communauté andine<br><i>Javier Herrera, Mireille Razafindrakoto et François Roubaud</i> ..... | 188 |
| 4.  | L'apport des outils décisionnels dans la réalisation des enquêtes des pays en<br>développement : le cas de Madagascar et du Mali<br><i>Jean-Marc Philip et Myriam Maumy</i> .....                                                                        | 194 |
| 5.  | Peut-on se fier aux bases de données internationales sur la corruption ? Une<br>confrontation entre enquêtes-experts et enquêtes-ménages en Afrique subsaharienne<br><i>Mireille Razafindrakoto et François Roubaud</i> .....                            | 199 |
| 6.  | Effet du plan de sondage dans les enquêtes 1-2-3 en Afrique de l'Ouest<br><i>Aude Vescovo et Eloi Ouedraogo</i> .....                                                                                                                                    | 204 |
| 7.  | Précision des indicateurs de niveau de vie de l'enquête ECAM 2<br><i>Justin Bem, Martin MBA et Ludovic Subran</i> .....                                                                                                                                  | 209 |
| 8.  | Prévalence de l'obésité en Algérie : résultats d'une enquête sur un échantillon national<br><i>Mohamed Lamine Atif, Salim Benkhedda, Kamel Merad-Boudia et<br/>Abdeldjalil Bezzaoucha</i> .....                                                          | 215 |
| 9.  | Sondage simplifié et coût-efficace dans les enquêtes Origine-Destination (OD)<br>de type entreprise<br><i>Omer Ramsés Zang Sidjou</i> .....                                                                                                              | 219 |
| 10. | Exemple de méthodologie pour une enquête auprès des populations<br>d'Afrique subsaharienne<br><i>Nathalie Lydié, Philippe Guilbert et Gaël Sliman</i> .....                                                                                              | 224 |



**CHAPITRE 6 • MÉTHODES DE COLLECTE ET APPLICATIONS**

|     |                                                                                                                                                                                                                       |     |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1.  | Techniques d'orientation et de manipulation dans les questionnaires<br><i>Gérard Dahan</i> .....                                                                                                                      | 229 |
| 2.  | Du questionnaire papier au support électronique : l'expérience du recensement canadien de 2006<br><i>Rossana Manriquez</i> .....                                                                                      | 234 |
| 3.  | Questionnaire : apports et limites de la rotation des items<br><i>Robert Revat</i> .....                                                                                                                              | 238 |
| 4.  | La technique du questionnaire unique<br><i>Souleymane Ouedraogo</i> .....                                                                                                                                             | 243 |
| 5.  | L'apport d'un échantillon de possesseurs exclusifs de téléphones mobiles dans l'enquête « Contexte de la sexualité en France » (2005-2006)<br><i>Laurent Toulémon, Nicolas Razafindratsima et le groupe CSF</i> ..... | 249 |
| 6.  | L'acharnement téléphonique dans les enquêtes est-il justifié ?<br><i>François Beck, Arnaud Gautier, Philippe Guilbert et Pierre Arwidson</i> .....                                                                    | 254 |
| 7.  | Enquête EFE : retour méthodologique sur la collecte (papier et Internet)<br><i>Céline Dauplait, Bernard de Clédât et Tania Vichnevskaja</i> .....                                                                     | 259 |
| 8.  | Utilisation de modes de recueil de données mixtes dans une enquête auprès des étudiants<br><i>Jean-Baptiste Combes, Valérie Guagliardo, Patrick Perreti-Watel et Pierre Verger</i> .....                              | 264 |
| 9.  | D'une approximation statistique à une approximation géographique : l'exemple d'une enquête santé à Luxembourg<br><i>Philippe Gerber et Sébastien Fleuret</i> .....                                                    | 269 |
| 10. | Enquête Web : une opportunité pour réduire la non-réponse totale dans les enquêtes ménages déplacements<br><i>Caroline Bayart et Patrick Bonnel</i> .....                                                             | 275 |
| 11. | Tests méthodologiques sur les questionnaires « en ligne »<br><i>Olivier Lê Van Truoc et Benoît Parraud</i> .....                                                                                                      | 279 |

**CHAPITRE 7 • ÉCHANTILLONNAGE**

|    |                                                                                                                                                                 |     |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1. | Contraintes d'équilibrage non linéaires<br><i>Éric Lesage</i> .....                                                                                             | 285 |
| 2. | Sondage systématique et sondages à support minimal<br><i>Johan Pea, Lionel Qualité et Yves Tillé</i> .....                                                      | 290 |
| 3. | Échantillonnage sous contraintes à entropie maximale : un algorithme rapide basé sur l'optimisation linéaire en nombres binaires<br><i>Philippe Périé</i> ..... | 294 |
| 4. | Amélioration de l'allocation de Neyman en tenant compte de la précision dans les strates de publication<br><i>Malik Koubi et Sandrine Mathern</i> .....         | 299 |
| 5. | Coordination des échantillons en utilisant la sélection contrôlée<br><i>Alina Matei</i> .....                                                                   | 305 |
| 6. | Tirages coordonnés d'échantillons stratifiés : méthodes basées sur des microstrates<br><i>Desislava Nedyalkova et Yves Tillé</i> .....                          | 309 |
| 7. | Échantillonnage sur les flux de données : état de l'art<br><i>Raja Chiky, Alain Dessertaine et Georges Hébrail</i> .....                                        | 314 |

**CHAPITRE 8 • ESTIMATION**

|     |                                                                                                                                                                                                                                   |     |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1.  | Calages successifs, calages itérés<br><i>Jean-Claude Deville</i> .....                                                                                                                                                            | 319 |
| 2.  | Sur l'utilisation de la régression non paramétrique pour des données d'enquêtes<br><i>Torsten Harms et Pierre Duchesne</i> .....                                                                                                  | 327 |
| 3.  | Un échantillon double pour minimiser le biais dû à la non-réponse dans<br>un sondage par courrier<br><i>Herbert L. Smith</i> .....                                                                                                | 334 |
| 4.  | Sondage à deux systèmes d'unités primaires couplées<br><i>Sébastien Hallépée et Vincent Loonis</i> .....                                                                                                                          | 339 |
| 5.  | Intégration de deux sources de données dans les inventaires forestiers du Québec<br><i>Sophie Baillargeon et Louis-Paul Rivest</i> .....                                                                                          | 344 |
| 6.  | Comparaison de trois méthodes d'estimation du loyer fictif<br><i>Modou Dia</i> .....                                                                                                                                              | 348 |
| 7.  | Estimation de courbes de consommation électrique à partir de mesures<br>non synchrones<br><i>Alain Dessertaine</i> .....                                                                                                          | 353 |
| 8.  | Méthodes de Bootstrap en population finie<br><i>Guillaume Chauvet</i> .....                                                                                                                                                       | 357 |
| 9.  | Une méthode Bootstrap pratique pour tester des hypothèses à partir<br>de données d'enquête<br><i>Jean-François Beaumont et Cynthia Bocci</i> .....                                                                                | 365 |
| 10. | Estimation de variance en présence de deux échantillons : linéarisation et Bootstrap<br><i>Guillaume Chauvet, Camélia Goga et Anne Ruiz-Gazen</i> .....                                                                           | 369 |
| 11. | Évaluation de précision par Bootstrap pour des algorithmes de tirage<br>à probabilités inégales : une étude empirique<br><i>Aurore Kervella, Emmanuel L'Hour, Nicolas Raillard, Stevonn Volant et<br/>Guillaume Chauvet</i> ..... | 374 |

**INDEX DES AUTEURS**



# Avant-propos

L'abondance toujours plus grande de données administratives pourrait laisser croire que le recours aux enquêtes par échantillonnage deviendra vite une technique obsolète. A quoi bon constituer des estimations basées sur des données partielles si des données exhaustives sont disponibles ? Pourtant, les données d'enquêtes sont de plus en plus utilisées. En témoigne le nouveau recensement français, où un relevé décennal a été remplacé par l'utilisation annuelle de groupes de rotation en appliquant les techniques d'échantillonnage et d'estimation les plus modernes. Ce nouveau recensement, rénové en partie grâce à l'utilisation de méthodes de sondage, permet maintenant une publication annuelle des statistiques démographiques.

Les statisticiens d'enquête doivent affronter de nouveaux défis. De plus en plus d'informations sont disponibles par voies administratives, mais les fichiers existants ne contiennent pas nécessairement les informations pertinentes. Les statisticiens doivent donc mettre à profit ces sources parfois hétéroclites pour concevoir les plans de sondage et pour estimer les paramètres d'intérêt. L'utilisation de sources multiples en statistique publique est un enjeu majeur qui nécessitera de nouveaux développements théoriques et méthodologiques.

Les huit chapitres de ce livre reprennent les principales communications du cinquième *colloque francophone sur les sondages* qui s'est tenu à Marseille en novembre 2007 sous l'égide du groupe *Enquêtes, modèles et applications* de la *Société française de statistique*. Ce livre est destiné aux chercheurs, aux praticiens d'enquêtes et aux enseignants qui veulent actualiser leurs connaissances et s'intéressent aux applications les plus récentes de la théorie des sondages. Cet ouvrage est le cinquième tome d'une série publiée par Dunod reprenant les interventions des quatre colloques précédents qui se sont

tenus à Rennes en 1997, Bruxelles en 2000, Grenoble en 2002 et Québec en 2005.

*Le colloque francophone sur les sondages* est un lieu d'échange important où sont exposées les dernières innovations statistiques de la théorie des sondages et les applications les plus récentes. En 2007, le comité scientifique a voulu organiser un colloque tourné vers les applications dans les domaines de la santé et des enquêtes électorales, en espérant que le dialogue entre la théorie et les applications soit fécond. L'analyse des enquêtes longitudinales a également été un thème privilégié, car c'est un outil crucial dans le domaine de la santé publique. Le comité scientifique a également souhaité que les chercheurs de tous les pays francophones aient la possibilité de participer à cette conférence. Dans cet esprit, les meilleures communications émanant de pays en développement ont été sélectionnées. La participation de leurs auteurs au colloque sondage a été financée par des bourses. Nous tenons à remercier la Fondation « la Science Statistique » présidée par le Professeur Edmond Malinvaud, qui a contribué au financement de ces bourses.

Le premier chapitre porte sur la qualité des données et la statistique publique. L'estimation à partir de sources multiples est l'un des thèmes privilégiés avec un premier article de Jean-Michel Charpin, ancien directeur général de l'Institut National de la Statistique et des Études Économiques (Insee), et deux contributions, illustrant d'une part la nécessité de mettre en œuvre une vérification sélective des données sur une enquête de l'Insee et cherchant d'autre part à expliquer les divergences des données des différentes sources sur une enquête de l'Institut national d'études démographiques (Ined). Ce chapitre contient aussi une description de la mise en place d'outils de gestion de la qualité à l'Institut de la Statistique du Québec ainsi qu'une présentation du nouvel échantillon maître de l'Insee. Les difficultés de la mise en œuvre d'enquêtes internationales font l'objet de deux articles avec l'exemple de quatre grandes enquêtes internationales en sciences sociales et l'étude de la qualité des données d'entreprises sur une enquête coordonnée par Eurostat.

Le thème des enquêtes longitudinales est abordé dans le chapitre 2. L'article de Gad Nathan et Abdulhakeem Eideh aborde la question du traitement de la non-réponse dans les enquêtes répétées. Michelle Simard décrit ensuite les méthodes d'imputation dans les enquêtes longitudinales au Canada. Puis, une contribution traite de la question du calcul de précision dans l'enquête emploi de l'Insee. Vient ensuite la présentation de deux enquêtes consacrées aux enfants en France : une enquête sur panel d'élèves entrés en 6<sup>ème</sup> en 1995 et la mise en place d'un système d'observation longitudinale des enfants en danger. S'ensuivent deux textes étudiant le phénomène de l'attrition : le premier dans le contexte du panel suisse de ménages et le second dans le contexte de l'enquête « Mobilité, attitudes, risque et comportements » (MARC) visant à étudier le risque routier des jeunes automobilistes français.

Le chapitre 3, consacré aux enquêtes dans le domaine de la santé, est introduit par Marcel Goldberg qui propose un tour d’horizon des principes méthodologiques et des difficultés rencontrées dans les cohortes épidémiologiques prospectives. D’autres enquêtes ont pour vocation d’estimer des prévalences ou de décrire des comportements de santé au niveau national et régional. Deux exemples d’enquêtes multithématiques de santé publique sont ici présentés. Il s’agit des enquêtes de Statistique Canada, de l’Office fédéral de la statistique (Suisse) et de l’Institut national de la prévention et de l’éducation pour la santé (Inpes, France). Le chapitre 3 donne également des exemples d’enquêtes thématiques avec le détail de leur méthodologie. Il s’agit d’une enquête sur la nutrition de l’Agence française de sécurité sanitaire des aliments (AFSSA), d’une enquête sur la prévalence de VIH et du VHC de l’Institut de veille sanitaire (InVS), d’une enquête auprès de personnes actuellement soignées présentée par Patrick Peretti-Watel et d’une étude sur l’opportunité d’interroger une personne en capacité de répondre (*répondant proxy*) à la place de celle sélectionnée par le protocole mais en incapacité de le faire. Deux autres articles rappellent l’importance du lien entre santé et événement de vie : une enquête de la Direction de la recherche, des études, de l’évaluation et des statistiques (DREES) et une de l’Université de Liège.

Les sondages électoraux font l’objet du chapitre 4 qui commence par une présentation et une analyse du rôle de la commission des sondages en France par Elisabeth Dupoirier. L’article suivant étudie l’impact du positionnement des questions sur les intentions de vote pour des enquêtes électorales canadiennes. Les trois articles suivants traitent des élections présidentielles françaises de 2007 et présentent une analyse de l’espace politique français et deux études sur l’indétermination et l’hésitation des électeurs dans les enquêtes préélectorales. Finalement, François Yale et Claire Durand analysent les effets de formulation de la question et les effets de contexte dans les enquêtes concernant la souveraineté du Québec.

Le chapitre 5, dédié aux sondages réalisés dans les pays en développement, offre une photographie des principaux enjeux et des principales difficultés de la statistique publique dans ces pays. Les responsables des enquêtes statistiques doivent souvent faire face à des pénuries de moyens humains et financiers. Najwa Yaacoub décrit ainsi toute la mise en place de l’Institut de Statistique du Liban à travers l’évolution de la situation politique et économique de son pays. Dans des contextes parfois délicats, les instituts nationaux de statistique ou les chercheurs doivent souvent faire preuve d’un grand sens pragmatique et jouer d’astuces pour limiter les coûts de collecte (sondage simplifié, recours aux outils décisionnels). Les sujets des enquêtes exposés dans le cadre de ce colloque montrent également à quel point la statistique des pays en développement peut jouer un rôle majeur d’alerte aussi bien auprès des pouvoirs publics que de l’ensemble de la communauté internationale (estimation du nombre d’enfants victimes de la traite au Bénin,

lutte contre la pauvreté, estimation de la corruption). Le lecteur trouvera aussi des sujets moins spécifiques à ces pays avec un article sur l'estimation de la précision dans une enquête « ménage » camerounaise, un autre sur l'estimation de l'obésité en Algérie ou encore une étude sur le rôle des enquêtes dans les prises de décisions politiques. Un dernier article présente les comportements, attitudes, opinions et connaissances des populations subsahariennes vivant en France dans le domaine du sida/VIH.

Les contributions du chapitre 6 portent sur les méthodes de collecte et applications. Les points successivement abordés sont les techniques d'orientation et de manipulation dans les questionnaires, l'analyse de nouveaux modes de saisie dans le recensement canadien de 2006, les apports et limites de la rotation des items dans un questionnaire, la technique du questionnaire unique, l'apport d'un échantillon de possesseurs exclusifs de téléphones mobiles sur une enquête de l'Institut national de la santé et de la recherche médicale (Inserm), l'étude de l'acharnement téléphonique, la description et l'évaluation d'un double mode de collecte, (internet et papier), pour une enquête de l'Ined, l'analyse de l'interchangeabilité de différentes stratégies de collecte sur une enquête de l'Observatoire régional de la santé, Provence-Alpes-Côte d'Azur (ORS PACA), la mise au point d'une reconstitution de répartition spatiale à partir de données échantillonnées dans la ville de Luxembourg, l'utilisation du web comme une opportunité pour réduire la non-réponse totale et enfin l'évaluation de l'incidence de la mise en forme des questions dans les enquêtes « en ligne ».

Au chapitre 7, le lecteur trouvera des développements formels récents sur plusieurs aspects de l'échantillonnage. Deux articles sont consacrés au problème de l'échantillonnage équilibré : un algorithme utilisant une technique de parcours 'intelligente' de l'ensemble des solutions dite de « Branch and Bound » et une méthode permettant d'équilibrer sur des contraintes non-linéaires. Deux textes traitent du problème de la coordination d'échantillons. L'un étudie une méthode de sélection contrôlée et l'application de cette méthode à la coordination des échantillons. L'autre propose huit nouvelles méthodes basées sur des permutations de nombre aléatoires dans des microstrates et compare ces dernières à des méthodes existantes. Les autres contributions portent sur les propriétés de l'échantillonnage systématique, sur le problème de la répartition de l'échantillon dans le contexte de l'échantillonnage stratifié et sur l'échantillonnage sur les flux de données.

Plusieurs questions relatives à l'estimation en populations finies font l'objet du chapitre 8. Jean-Claude Deville développe les propriétés théoriques du calage successif. Pierre Duchesne présente les propriétés asymptotiques des estimateurs polynomiaux locaux faisant intervenir le plan de sondage. Guillaume Chauvet fait une synthèse des méthodes bootstrap applicables en population finie. L'estimation de la variance par la méthode du bootstrap en

populations finies est traitée dans deux autres textes. Une méthode utilisant les poids bootstrap pour tester des hypothèses sur des paramètres analytiques est présentée. Ce chapitre contient aussi la présentation de l'utilisation du modèle linéaire mixte pour l'intégration de deux sources de données dans les inventaires forestiers du Québec, une comparaison de trois méthodes d'estimation de loyer fictif, l'utilisation de l'échantillonnage à deux-phases pour réduire les biais de non-réponse, une présentation de certaines facettes de l'estimation dans le cadre de l'enquête Vie Quotidienne et Santé ainsi qu'une description de l'estimation de courbes de consommation électrique à partir de mesures non synchrones.

Nous voulons remercier toutes les personnes qui ont contribué au succès du cinquième *colloque francophone sur les sondages* en particulier les membres du comité scientifique<sup>1</sup> et du comité d'organisation<sup>2</sup> qui ont consacré beaucoup de temps pour que cette manifestation soit une réussite. Les principaux partenaires de cette manifestation ont été l'Observatoire Régional de la Santé (ORS) de Provence-Alpes-Côte d'Azur, la Direction régionale Provence Alpes Côte d'Azur de Institut national de la statistique et des études économiques (Insee), le Centre d'études et de recherches sur les qualifications (Céreq), le Laboratoire population-environnement-développement (Lped), le Laboratoire d'économie et de sociologie du travail (Lest), l'Institut fédératif de recherches sciences humaines économiques & sociales d'Aix-Marseille (IFR 134 SHESS-AM) et le Centre sur la Population et le Développement (CEPED). Nous exprimons également notre gratitude envers les institutions et les entreprises suivantes qui, par leur soutien financier, ont permis la réalisation du colloque et de cet ouvrage : Ined, Insee, Fondation pour la science, Cnrs, Drees, InVS, Inpes, Office fédéral suisse de la statistique, Institut de la statistique du Québec, Ville de Marseille, BVA, IPSOS, LH2, IRDES. Nous remercions les auteurs des articles de cet ouvrage qui ont travaillé avec diligence pour se conformer à nos exigences éditoriales. Enfin, nous voulons remercier tout particulièrement Céline Ethier qui a réalisé la

---

1 Yves Tillé (Président, Université de Neuchâtel, Suisse), Louise Bourque (Vice-Présidente, Institut de la statistique du Québec, Canada), Jean Chiche (CEVIPOF, CNRS, Sciences Po, Paris), Vanessa Di Paola (LEST, Université de la Méditerranée, Aix-en-Provence), Philippe Guilbert (INPES, Paris), David Haziza (Université de Montréal, Montréal), Farhad Mehran (Ancien directeur du service statistique du BIT, Genève), Jean-Paul Moatti (INSERM U379, Marseille), Camelia Protopopescu, (INSERM U379, ORS PACA, Marseille), Anne Ruiz-Gazen (Université Toulouse 1), Paul-André Salamin (Office Fédéral de la Statistique, Neuchâtel), Olivier Sautory (Insee, Paris), Laurent Toulemon (INED, Paris).

2 Docteur Yolande Obadia, Présidente, directrice de l'Observatoire régional de santé PACA, François Clanché, vice-président et trésorier, directeur régional de l'Insee PACA, Odile Bellardi (INSEE PACA), Vanessa Di Paola (LEST, Université de Provence, Aix en Provence), Alberto Lopez (Céreq, Marseille), William Molmy (Céped, Paris), Jean-Luc Perrot (Insee PACA), Benoît Riandey (SFdS, INED, Paris), Michel Stoësz (Céreq, Marseille), Jacques Vaugelade (LPED, IRD, Marseille), Christine Voto (ORS, PACA).



mise en page de ce livre et Soazig L'Helgouac'h qui a assuré la plus grande partie du travail de secrétariat du comité scientifique et du comité éditorial de cet ouvrage.

Philippe Guilbert, David Haziza, Anne Ruiz-Gazen et Yves Tillé

# Chapitre 1

## **Qualité des enquêtes et statistique publique**

### **1. Remarques sur l'estimation à partir de sources multiples**

*Jean-Michel CHARPIN*<sup>1</sup>

#### **1.1 Introduction**

Après presque cinq années à la tête de l'Insee, j'ai acquis une expérience pratique de l'estimation statistique à partir de sources multiples. C'est sur la base de cette expérience que je formulerai quelques remarques modestes. Même quand je pourrais le faire, je ne chercherai pas à entrer dans les aspects scientifiques de l'estimation à partir de sources multiples.

#### **1.2 Première remarque : la comptabilité nationale nous a habitués à procéder de façon courante à des estimations à partir de sources multiples**

Dans la comptabilité nationale, il y a un chiffre et un seul dans chaque case.

Il y a un chiffre. Il faut donc en mettre un, même quand on n'a pas d'informations. L'estimation se fait alors par solde ou en appréciant la

---

<sup>1</sup> Inspecteur général des Finances, ancien directeur général de l'Insee de janvier 2003 à octobre 2007.

cohérence en ligne et en colonne. Cette méthode, qui surprend toujours les vrais comptables, permet à la comptabilité nationale d'être disponible très rapidement. Ce fut l'une des raisons qui firent choisir la comptabilité nationale comme langage de la procédure de déficit excessif créée par le Traité de Maastricht, décision qui n'allait pas de soi.

Mais surtout, il n'y a qu'un seul chiffre. Lorsqu'existent des sources multiples, ce qui est fréquent, il est alors nécessaire d'en choisir un parmi tous les candidats possibles, de procéder à des arbitrages entre les différentes sources.

Cette contrainte d'unicité, de synthèse est très exigeante. En même temps, elle a sa part dans l'explication du succès considérable, peut-être excessif, de la comptabilité nationale tant dans le débat public que dans les procédures politiques, particulièrement au niveau européen.

Dans la plupart des pays, ce sont les statisticiens qui sont responsables de l'établissement de la comptabilité nationale. Ils ont donc acquis une expérience de l'estimation à partir de sources multiples. Le principal enseignement qu'ils en ont tiré est qu'il faut éviter de trop fréquents changements dans les procédures d'arbitrage. Ce sont les variations des agrégats comptables qui intéressent au premier chef les utilisateurs, plus que la précision des estimations en niveau. Or les changements ont un impact direct sur les variations. La fidélité de l'instrument est donc essentielle. Elle nécessite beaucoup de stabilité, voire de conservatisme : les changements ne doivent intervenir qu'en cas d'apparition de sources nouvelles ou de fort basculement dans la qualité respective des sources.

### **1.3 Deuxième remarque : de plus en plus, la statistique elle-même estime à partir de sources multiples**

Je vais en donner deux exemples qui me paraissent très illustratifs.

#### **1.3.1 Les populations légales en France**

La loi du 27 février 2002, qui a réformé le recensement français, prévoit que l'Insee établit les populations légales en utilisant les enquêtes de recensement, certes, mais aussi les données issues des fichiers administratifs ainsi que de toutes les enquêtes statistiques. À lire l'article concerné de cette loi, on comprend clairement que, sur le plan juridique, ce n'est plus le recensement qui détermine la population légale, c'est la décision de l'Insee, fondée sur l'ensemble des sources à sa disposition.

Le hasard a d'ailleurs fait que nous avons très rapidement fait usage de cette possibilité lors du recensement conduit en 2004 en Nouvelle-Calédonie, alors même qu'il s'agissait d'une enquête de recensement exhaustive, à l'ancienne, mais qui avait donné lieu à des mouvements localisés de boycott.