

Techniques d'enquête

Commentaires à propos de l'article « Inférence statistique avec des échantillons d'enquête non probabiliste »

par Sharon L. Lohr

Date de diffusion : le 15 décembre 2022



Statistique
Canada

Statistics
Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par le ministre de l'Industrie 2022

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Commentaires à propos de l'article « Inférence statistique avec des échantillons d'enquête non probabiliste »

Sharon L. Lohr¹

Résumé

Des hypothèses solides sont nécessaires pour faire des inférences au sujet d'une population finie à partir d'un échantillon non probabiliste. Les statistiques d'un échantillon non probabiliste devraient être accompagnées de preuves que les hypothèses sont respectées et que les estimations ponctuelles et les intervalles de confiance sont propres à l'utilisation. Je décris certains diagnostics qui peuvent être utilisés pour évaluer les hypothèses du modèle, et je discute des questions à prendre en considération au moment de décider s'il convient d'utiliser les données d'un échantillon non probabiliste.

Mots-clés : Échantillon de commodité; diagnostic; imputation; échantillon probabiliste; qualité de l'enquête; poids d'enquête.

1. Introduction

Un grand merci à Changbao Wu pour son examen stimulant et son évaluation des méthodes permettant de faire des inférences à partir d'échantillons non probabilistes. J'apprécie particulièrement son examen réfléchi des hypothèses solides nécessaires pour établir le biais et la variance des estimations.

M. Wu passe en revue trois approches pour estimer la moyenne de population finie μ_y d'une variable y qui est mesurée dans un échantillon S_A non probabiliste de taille n_A . Étant donné que cet échantillon n'est pas représentatif de la population (et que, par conséquent, la moyenne de l'échantillon \bar{y}_A est probablement biaisée pour estimer μ_y), chaque approche repose sur l'information provenant d'un échantillon probabiliste S_B de grande qualité et de taille n_B : S_B ne mesure pas y , mais contient un ensemble de variables auxiliaires \mathbf{x} qui sont également observées dans S_A .

Dans l'approche prédictive basée sur un modèle, un modèle est créé sur S_A pour prédire y à partir de \mathbf{x} . L'estimateur d'imputation massive (IM), par exemple, utilise le modèle pour imputer une estimation y_i^* de y_i pour chaque membre de l'échantillon probabiliste S_B . Ensuite, le total de la population de y est estimé par $\sum_{i \in S_B} d_i^B y_i^*$ où d_i^B est le poids d'enquête de l'unité i dans S_B .

Dans l'approche de la pondération par l'inverse de la propension (PIP), un modèle est élaboré pour prédire la probabilité π_i^A que l'unité de population i apparaisse dans S_A en fonction de \mathbf{x} . L'unité i dans S_A reçoit ensuite la pondération $w_i^A = 1/\hat{\pi}_i^A$ et le total de la population est estimé par $\sum_{i \in S_A} w_i^A y_i$.

M. Wu se penche également sur un estimateur « doublement robuste » de μ_y qui, en combinant les estimateurs prédictifs et de PIP, est approximativement sans biais selon les hypothèses que l'un ou l'autre des modèles est correctement indiqué. Dans cet exposé, je me concentrerai sur les approches prédictives et de PIP, parce que ces méthodes se généralisent plus facilement pour les analyses multivariées et l'estimation des caractéristiques des populations autres que les moyennes.

1. Sharon Lohr est professeure émérite à l'Arizona State University. Courriel : sharon.lohr@asu.edu.

À la section 2, j'examinerai les hypothèses nécessaires à l'inférence à partir d'échantillons non probabilistes et des diagnostics pour les évaluer. Ensuite, à la section 3, je me pencherai sur certaines questions à poser au moment de décider quelle approche (le cas échéant) utiliser pour l'inférence.

2. Hypothèses et diagnostics de modèle

L'échantillonnage probabiliste a été largement utilisé après l'élaboration de la théorie dans les années 1930 et 1940 parce qu'il offrait une solution mathématiquement justifiée au problème de la généralisation d'un échantillon à une population. Selon des hypothèses minimales, un échantillon probabiliste avec réponse complète produit des estimations approximativement sans biais des quantités de population, accompagnées d'intervalles de confiance qui ont des probabilités de couverture approximativement exactes. C'est la *seule* méthode qui garantit la création d'intervalles de confiance exacts sans qu'il soit nécessaire de poser d'hypothèses sur les membres non échantillonnés d'une population donnée. Un échantillon probabiliste est représentatif en raison de la procédure utilisée pour l'établir.

Toutes les autres méthodes exigent d'énormes hypothèses. Les principales hypothèses applicables aux méthodes prédictives et de PIP, qui sont présentées à la section 2.1 de l'article de M. Wu, sont les suivantes : (A1) y et la variable aléatoire indiquant la participation à S_A sont indépendants étant donné \mathbf{x} ; (A2) chaque unité de la population se caractérise par $\pi_i^A > 0$, et (A3) les variables aléatoires indiquant la participation à S_A sont indépendantes étant donnée \mathbf{x} . Ces hypothèses laissent entendre que l'information auxiliaire \mathbf{x} est suffisamment riche pour élaborer des poids de propension inverse qui éliminent le biais de sélection pour y , et qu'un modèle élaboré sur S_A pour prédire y à partir de \mathbf{x} s'appliquera également aux unités qui ne sont pas dans S_A .

Les propriétés statistiques des estimateurs sont élaborées en supposant que les hypothèses (A1) à (A3) sont vraies et que les modèles adoptés pour la pondération ou l'imputation sont correctement indiqués. Dans ces conditions, la moyenne estimée de la population est approximativement sans biais avec la variance obtenue par le théorème approprié. Mais, comme M. Wu le souligne dans son article, cette estimation de la variance est conditionnelle au respect des hypothèses. Autrement, l'erreur quadratique moyenne réelle sera sérieusement sous-estimée et la fiabilité de l'estimation donnera une impression trompeuse. Si n_A et n_B sont importants, mais que (A1) n'est pas respectée, le biais pourrait être de 10 points de pourcentage, mais l'erreur-type déclarée d'une estimation d'IM ou de PIP sera proche de zéro. Dans la pratique, de nombreux échantillons non probabilistes ne respecteront pas les hypothèses. Mercer, Lau et Kennedy (2018) ont constaté, en pondérant les échantillons d'enquêtes volontaires en ligne avec de l'information auxiliaire riche, que « même la stratégie de rajustement la plus efficace n'a permis d'éliminer qu'environ 30 % du biais initial ».

Il est impossible de mettre les hypothèses pleinement à l'essai parce qu'elles se caractérisent pas des données manquantes : des membres de la population sont absents de S_A , et des valeurs y sont absentes de S_B . Toutefois, comme pour les rajustements effectués pour compenser la non-réponse dans les

échantillons probabilistes (Lohr, 2022, chapitre 8), il est possible d'effectuer des vérifications et des diagnostics de modèle à l'aide des renseignements disponibles, tout en étant conscient que cela risque de ne pas déceler toutes les lacunes du modèle.

Comparaison des statistiques de l'échantillon non probabiliste avec celles d'autres sources de données

M. Wu suggère de comparer les fonctions de distribution empirique des variables contenues dans \mathbf{x} à partir de S_A avec les fonctions de distribution empirique pondérées par des poids d'enquête tirés de S_B . Les différences peuvent indiquer que les observations dans S_A ont des scores de propension inégaux, ou que les variables \mathbf{x} dans S_A et dans S_B sont mesurées de façon distincte (voir la section 3). Il est également possible de comparer les distributions empiriques de S_A avec celles de S_C d'une autre enquête probabiliste.

Si l'on utilise la PIP, on peut aussi comparer les fonctions de distribution empirique pondérées par le score de propension de S_A avec celles de S_B et d'autres enquêtes. Cela ne devrait être fait que pour les variables qui ne sont pas utilisées dans la pondération, puisque les poids des scores de propension ont déjà été rajustés pour tenir compte des déséquilibres dans les variables de pondération. Dutwin et Buskirk (2017), par exemple, ont construit des poids de propension pour un échantillon non probabiliste en ratissant des totaux marginaux, puis en comparant les tableaux croisés de ces variables de ratissage.

M. Wu suggère également de traiter une variable z qui est mesurée à la fois dans S_A et S_B en tant que variable réponse, et de comparer les modèles conditionnels de $z|\mathbf{u}$ ajustés sur S_A et S_B , où \mathbf{u} est un sous-ensemble de \mathbf{x} (à l'exclusion de z). Les différences entre les deux modèles peuvent indiquer que z est nécessaire en tant que variable auxiliaire, et peuvent également soulever la question de savoir dans quelle mesure l'ensemble de variables auxiliaires mesurées répond à l'hypothèse (A1).

Dans un exemple de Kim, Park, Chen et Wu (2021), le pourcentage estimé de personnes qui s'étaient portées volontaires était de 24,8 % selon la *Current Population Survey* (estimation de référence), mais les estimations de l'IM et de la PIP tirées de S_A étaient toutes deux proches de 50 % avec une erreur-type déclarée de moins d'un point de pourcentage. L'erreur-type, calculée selon les hypothèses du modèle, n'a pas tenu compte du biais de sélection de S_A en lien avec la participation volontaire, un biais qui ne pouvait pas être éliminé en utilisant les données démographiques, l'accession à la propriété et l'assurance médicale comme covariables du modèle.

Comparer les résultats des approches de la PIP et de l'IM

Une solution de rechange à l'utilisation d'un estimateur doublement robuste aux fins d'analyse consiste à utiliser chaque modèle pour déterminer les lacunes possibles de l'autre. Les examens possibles consistent à comparer la distribution empirique de y tirée de S_A (à l'aide des poids de propension inverse) avec la distribution empirique de y^* tirée de S_B (en utilisant les valeurs imputées et les poids d'enquête). De même, comme l'ont suggéré Chipperfield, Chessman et Lim (2012), on peut comparer les moyennes de domaine estimées à partir de S_A et S_B pour un ensemble de domaines $d=1, \dots, D$. On

pourrait aussi comparer les imputations pour l'ajustement de y à l'ensemble de données non pondérées S_A avec des imputations élaborées pour S_A avec des poids de propension inverse.

Les études de simulation sont utiles pour vérifier le comportement d'un petit échantillon lorsque les hypothèses sont respectées, mais elles ont un intérêt limité pour étudier la sensibilité aux hypothèses du modèle. Ces études permettent d'analyser les écarts entre les modèles conçus par les chercheurs, mais les enquêtes réelles peuvent diverger du modèle de nombreuses façons imprévues.

Effectuer des diagnostics de modèle

Bien entendu, qu'il s'agisse de la PIP ou de l'approche prédictive basée sur un modèle, les analystes devraient utiliser des diagnostics de régression standard, comme l'examen des résidus et les observations influentes pour étudier l'ajustement du modèle et la sensibilité aux valeurs aberrantes, et garder une trace des vérifications ainsi effectuées.

Dans le cas de l'approche de PIP, il est également souhaitable d'examiner les caractéristiques des poids finaux. Le coefficient de variation des poids fournit une mesure approximative de l'ampleur des ajustements qui ont été nécessaires pour rendre l'échantillon S_A représentatif. Toutefois, un faible coefficient de variation ne signifie pas nécessairement que l'échantillon est représentatif; cela peut simplement refléter l'insuffisance d'information auxiliaire pour élaborer des poids. Par exemple, supposons qu'un échantillon par quota d'un groupe à participation volontaire sur Internet est extrait pour le faire correspondre à la population en ce qui concerne les variables auxiliaires. Les poids de propension inverse connaîtront peu de variation parce que les variables \mathbf{x} ont été utilisées pour former les catégories de quotas, mais l'échantillon peut tout de même produire des estimations biaisées des variables y , comme l'utilisation d'Internet ou la participation volontaire.

Les méthodes graphiques proposées par Makela, Si et Gelman (2014) pour évaluer les ajustements de poids dans les enquêtes peuvent également être utilisées avec la PIP. Brick (2015) suggère d'examiner l'ampleur des ajustements de la PIP dans les classes de pondération. On peut également examiner la répartition des poids dans les domaines d'intérêt.

Les poids de propension inverse peuvent également fournir des renseignements sur l'hypothèse (A2). Un domaine qui a des pondérations élevées par rapport à d'autres domaines peut signaler un sous-dénombrement dans S_A . Dever (2018) a proposé d'examiner l'hypothèse (A2) en identifiant les répondants dans S_B qui n'ont pas de correspondance étroite dans S_A .

Bondarenko et Raghunathan (2016) ont examiné et proposé des outils de diagnostic graphiques et numériques pour évaluer et améliorer les modèles d'imputation. Cependant, aucun de ces diagnostics ne permet de valider l'hypothèse selon laquelle l'ajustement du modèle de régression sur S_A s'applique aux unités qui ne sont pas dans S_A . Tout comme \bar{y}_A pourrait être un estimateur biaisé de μ_y , les coefficients de régression dérivés de S_A peuvent également être biaisés, et le modèle construit sur S_A pour prédire y à partir de \mathbf{x} pourrait ne pas s'appliquer à d'autres parties de la population.

Prendre un petit échantillon probabiliste pour analyser les hypothèses

Les étapes précédentes peuvent permettre de déceler certaines lacunes du modèle, mais ne peuvent pas valider entièrement les hypothèses (A1) et (A2). Il est néanmoins possible de mettre à l'essai le modèle d'imputation en obtenant des données sur y à partir d'un sous-échantillon probabiliste de S_B . De même, on pourrait prendre un échantillon probabiliste de membres de la population qui ne sont pas présents dans S_A pour vérifier les inférences de l'approche de PIP, ou observer y sur un sous-échantillon d'unités dans S_B qui sont semblables à celles assorties d'une pondération élevée dans S_A , ou qui n'ont pas de correspondance étroite dans S_A .

3. Quand faut-il utiliser des échantillons non probabilistes ?

M. Wu décrit les méthodes pour combiner l'information des échantillons probabilistes avec celle des échantillons non probabilistes après qu'une décision a été prise à cette fin. Or, une première question est de savoir s'il faut réellement effectuer cette opération. On pourrait vouloir utiliser un échantillon non probabiliste parce qu'il n'existe pas d'échantillon probabiliste de grande qualité pour mesurer y , et qu'on pense que toute information est préférable à un manque d'information. Mais est-ce vrai ?

Supposons que, malgré l'ajustement minutieux et la vérification du modèle, les statistiques clés demeurent biaisées. La déclaration d'une statistique erronée pourrait-elle être pire que la déclaration d'aucune statistique ? Les mauvaises statistiques, une fois publiées, peuvent circuler longtemps, même après que des études plus rigoureuses montrent qu'elles sont biaisées. En 1975, la chroniqueuse Ann Landers a demandé à ses lecteurs de répondre à la question suivante : « Si c'était à refaire, auriez-vous des enfants ? » Environ 70 % des 10 000 personnes qui ont envoyé une réponse par la poste ont déclaré qu'elles n'auraient pas eu d'enfants si elles en avaient le choix. Cette statistique est toujours citée presque 50 ans plus tard, même si elle provient d'un échantillon de commodité et qu'elle a été contredite par de nombreuses autres études (Lohr, 2022). Il est également peu probable que la modélisation prédictive ou la PIP aurait corrigé le biais de sélection touchant la statistique de Landers, qui porte sur tous les groupes démographiques.

En gardant ces considérations à l'esprit, voici quelques questions qu'il conviendrait de poser au moment de décider d'utiliser des estimations d'un échantillon non probabiliste et, le cas échéant, quelle méthode statistique il faudrait adopter pour faire des inférences.

- Comment les statistiques seront-elles utilisées ? Les estimations d'un échantillon non probabiliste peuvent être utiles à l'élaboration d'une stratégie de marketing ou pour une étude sociologique exploratoire, mais elles pourraient ne pas être jugées suffisamment fiables pour estimer le chômage ou le nombre de personnes ayant besoin d'aide alimentaire. Les statistiques d'un échantillon non probabiliste devraient être accompagnées de preuves que les hypothèses sont propres à l'utilisation.

- Quelle est la qualité des données dans S_A ? Les dossiers administratifs, comme les dossiers fiscaux, ont un profil de qualité différent de celui d'une enquête sur des volontaires recrutés au moyen d'une annonce sur Internet.

Si la population de S_A est bien définie (il s'agit par exemple de déclarants fiscaux), il peut être préférable de déclarer des statistiques pour cette population plutôt que d'essayer de généraliser à la population de S_B . En ce qui concerne les dossiers fiscaux, de nombreuses personnes dont le revenu est inférieur aux seuils de revenu préétablis se caractérisent par la valeur $\pi_i^A = 0$, ce qui enfreint l'hypothèse (A2). On pourrait plutôt adopter une approche à bases multiples, où une source de données différente est utilisée afin d'estimer μ_y pour les parties de la population qui ne sont pas dans S_A (Lohr, 2021).

Étant donné que tous les modèles reposent sur des données auxiliaires \mathbf{x} , il est important que S_A et S_B mesurent les variables \mathbf{x} de la même façon. Si le revenu est utilisé comme variable auxiliaire, les mêmes questions doivent être utilisées pour définir le revenu dans les deux enquêtes, et le revenu doit être mesuré pour la même unité (personne ou ménage).

Kennedy (2022) indique que certains répondants qui choisissent de participer à des sondages en ligne peuvent fournir des renseignements démographiques inexacts ou de fausses réponses aux questions et que si cela se produit, les prévisions du modèle seront erronées. Il est même possible que des intervenants externes souhaitant obtenir un résultat particulier manipulent les données contenues dans S_A . Par exemple, une organisation pourrait s'arranger pour qu'une enquête soit attribuée à un ensemble de volontaires dont les caractéristiques démographiques prétendues correspondent à celles de la population, mais qui fournissent la réponse « souhaitée » pour y . Certains partisans des échantillons non probabilistes soutiennent que les échantillons probabilistes à faible taux de réponse nécessitent également des ajustements de poids ou une imputation, mais il y a une différence importante, à savoir que l'enquête probabiliste peut se caractériser par des non-réponses, mais l'échantillon initial est sélectionné au hasard et ne peut pas être manipulé par des organisations externes.

Si les données contenues dans S_A sont de faible qualité, vaut-il la peine de construire des modèles ? Comme l'a dit Louis (2016), « les procédures de l'ère spatiale ne sauveront pas les données de l'âge de pierre ».

- Dans quelle mesure les renseignements auxiliaires sont-ils détaillés ? Si S_A est importante et que l'information auxiliaire est suffisamment précise pour permettre de repérer des enregistrements précis, alors le couplage des enregistrements entre S_A et S_B serait une meilleure méthode pour combiner les données. On se servira de l'imputation ou de la PIP si l'information auxiliaire est suffisamment riche pour donner de bonnes prédictions de y_i ou de π_i^A , mais pas assez riche pour permettre un couplage précis. Cependant, s'il y a peu d'information auxiliaire, on peut s'attendre à une faible variation des scores de propension ou

des valeurs imputées, et les méthodes peuvent donner de mauvaises prévisions, étant donné le peu d'information permettant de diagnostiquer des problèmes éventuels.

- Quelles sont les analyses souhaitées ? M. Wu discute de l'estimation de la moyenne de la population, mais l'analyste pourrait aussi vouloir examiner les relations entre y et d'autres variables, ou estimer les moyennes ou les médianes pour les sous-groupes. Le choix de la méthode dépend en partie des variables disponibles dans S_A et S_B . Si S_A contient de nombreuses variables de réponse dont la relation présente un intérêt, l'approche de PIP pourrait être privilégiée.

Si l'on souhaite explorer les relations entre y et les variables mesurées uniquement dans S_B , l'imputation pourrait être un meilleur choix. Dans ce cas, cependant, l'analyste doit prendre soin de reconnaître l'imputation lorsqu'il présente les résultats. Si, par exemple, la régression linéaire est utilisée pour l'imputation, la corrélation calculée sur S_B n'est pas entre les variables u et y , mais entre u et $\mathbf{x}^T \hat{\boldsymbol{\beta}}$.

- Quelles sont les répercussions sur l'équité des données ? Jagadish, Stoyanovich et Howe (2021) ont défini l'« équité en matière de représentation » comme une « augmentation de la visibilité des groupes sous-représentés qui ont été historiquement désavantagés ou supprimés dans l'enregistrement de données ».

Les échantillons non probabilistes peuvent améliorer l'équité des données. Ils peuvent accroître la taille de l'échantillon et la visibilité des sous-groupes de population rares – un grand ensemble de données S_A pourrait contenir 10 000 membres d'un sous-groupe, tandis que même une enquête probabiliste à réponse complète de $n_B = 60\,000$ pourrait en contenir seulement 10. Ou encore, l'échantillon non probabiliste pourrait contenir des membres de la population qui sont sous-représentés dans l'enquête probabiliste parce qu'ils sont hors du champ de l'enquête, sous-dénombrés dans la base de sondage ou qu'ils présentent une tendance à la non-réponse. Dans ces situations, S_A fournit des renseignements sur les groupes qui ne sont pas aussi bien représentés dans l'enquête probabiliste.

Par ailleurs, les groupes historiquement défavorisés pourraient être sous-représentés dans toutes les sources de données, y compris dans S_A . Par exemple, un grand échantillon non probabiliste de dossiers de santé électroniques pourra produire des estimations pour un plus grand nombre de sous-groupes de population qu'un petit échantillon probabiliste sur la santé. Or, les personnes sans assurance-maladie ou qui n'ont pas accès aux soins médicaux seront sous-représentées. Dans cette situation, le fait d'avoir recours à S_A pour la production d'estimations démographiques est susceptible de renforcer les inégalités. Si des estimations sont utilisées pour répartir les ressources, alors, au fur et à mesure de la mise en œuvre du programme, davantage de données seront recueillies dans les régions qui obtiennent ces ressources et valideront leurs besoins, alors qu'aucun suivi ne sera effectué dans les régions qui sont désignées à tort comme

ne recevant aucune ressource. La boucle de rétroaction propagera la représentation inéquitable dans les sources de données.

Les méthodes d'IM et de PIP ont des répercussions différentes sur l'équité des données. L'imputation attribue une valeur prédite de y à chaque observation dans S_B , et la valeur imputée y peut différer de la valeur y que le répondant aurait fournie si on lui posait la question, en particulier si le répondant fait partie d'un sous-groupe qui est sous-représenté ou mal représenté dans S_A . Le modèle fournira-t-il des prédictions exactes pour les sous-groupes historiquement sous-représentés ? Les répondants de S_B ont-ils donné leur consentement éclairé pour permettre l'imputation de y ?

La PIP part du principe que les scores de propension peuvent être estimés à partir de renseignements auxiliaires. Cette information est-elle suffisamment riche pour donner des poids exacts ? Certains sous-groupes sont-ils sous-représentés dans S_A ? Il peut être utile de comparer les résultats des deux méthodes et d'autres sources de données, si elles sont disponibles, pour les sous-groupes de population historiquement sous-représentés.

L'examen critique de M. Wu soulève de nombreuses questions importantes pour les personnes qui envisagent d'utiliser des échantillons non probabilistes pour faire des inférences au sujet de la population. J'apprécie particulièrement son évaluation des hypothèses solides nécessaires pour les méthodes fondées sur des modèles, et je salue l'accent mis sur la résolution de ces problèmes à l'étape des plans d'enquête.

Bibliographie

- Bondarenko, I., et Raghunathan, T. (2016). Graphical and numerical diagnostic tools to assess suitability of multiple imputations and imputation models. *Statistics in Medicine*, 35(17), 3007-3020.
- Brick, J.M. (2015). Compositional model inference. Dans *Proceedings of the Survey Research Methods Section*, Alexandrie, Virginie: American Statistical Association, 299-307.
- Chipperfield, J., Chessman, J. et Lim, R. (2012). Combining household surveys using mass imputation to estimate population totals. *Australian & New Zealand Journal of Statistics*, 54, 223-238.
- Dever, J.A. (2018). Combining probability and nonprobability samples to form efficient hybrid estimates: An evaluation of the common support assumption. Dans *Proceedings of the 2018 Federal Committee on Statistical Methodology (FCSM) Research Conference*. https://nces.ed.gov/FCSM/pdf/A4_Dever_2018FCSM.pdf.
- Dutwin, D., et Buskirk, T.D. (2017). Apples to oranges or gala versus golden delicious? Comparing data quality of nonprobability internet samples to low response rate probability samples. *Public Opinion Quarterly*, 81(S1), 213-239.

- Jagadish, H.V., Stoyanovich, J. et Howe, B. (2021). COVID–19 brings data equity challenges to the fore. *Digital Government: Research and Practice*, 2(2), 1-7.
- Kennedy, C. (2022). Exploring the assumption that online opt-in respondents are answering in good faith. Communication présentée à la conférence 2022 Morris Hansen, 1^{er} mars 2022.
- Kim, J.-K., Park, S., Chen, Y. et Wu, C. (2021). Combining non-probability and probability survey samples through mass imputation. *Journal of the Royal Statistical Society, Series A*, 184, 941-963.
- Lohr, S.L. (2021). [Les enquêtes à bases de sondage multiples pour un monde fait de sources de données multiples](#). *Techniques d'enquête*, 47,2, 247-285. Article accessible à l'adresse <https://www150.statcan.gc.ca/n1/fr/pub/12-001-x/2021002/article/00008-fra.pdf>.
- Lohr, S.L. (2022). *Sampling: Design and Analysis, Third Edition*. Boca Raton, FL: CRC Press.
- Louis, T.A. (2016). Discussion of combining information from survey and non-survey data sources: Challenges and opportunities. Séminaire public de la 130^e réunion du CNSTAT; Washington, DC. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_172505.pdf.
- Makela, S., Si, Y. et Gelman, A. (2014). Statistical graphics for survey weights. *Revista Colombiana de Estadística*, 37(2), 285-295.
- Mercer, A., Lau, A. et Kennedy, C. (2018). *For Weighting Online Opt-In Samples, What Matters Most?* Washington, DC: PewResearch.