

Explorations de l'échantillonnage non probabiliste par Internet

J. Michael Brick¹

Résumé

Même si l'estimation de caractéristiques de populations finies à partir d'échantillons probabilistes a obtenu beaucoup de succès pour les grands échantillons, des inférences à partir d'échantillons non probabilistes sont également possibles. Les échantillons non probabilistes ont été critiqués en raison du biais d'auto-sélection et de l'absence de méthodes pour estimer la précision des estimations. L'accès plus vaste à Internet et la capacité de procéder à des collectes de données très peu coûteuses en ligne ont ravivé l'intérêt pour ce sujet. Nous passons en revue des stratégies d'échantillonnage non probabiliste et nous résumons certains des enjeux clés. Nous proposons ensuite des conditions à respecter pour que l'échantillonnage non probabiliste puisse constituer une approche raisonnable. Nous concluons par des idées de recherches futures.

Mots clés : Inférence, représentativité, biais d'auto-sélection

1. Introduction

L'échantillonnage probabiliste est généralement reconnu comme la méthode la plus appropriée pour procéder à des inférences qui peuvent être généralisées à une population finie. Cette méthode a une longue tradition et un fondement théorique solide, qui s'est révélé efficace dans de nombreuses études empiriques. Dans un échantillon probabiliste, chaque unité de la population comporte une probabilité connue non nulle d'être échantillonnée, et dans le cadre fondé sur le plan de sondage, ces probabilités constituent la base des inférences (Hansen, Hurwitz et Madow, 1953; Särndal, Swensson et Wretman, 1992; Lohr, 2009). Presque toutes les statistiques officielles reposent sur cette méthodologie, et de nombreux instituts nationaux de statistique ont recours à l'échantillonnage probabiliste pour faire des inférences.

Toutefois, l'échantillonnage probabiliste n'est pas la seule méthode qui permet de tirer des échantillons et de faire des inférences. En fait, au cours du XX^e siècle, la transition vers l'échantillonnage probabiliste a commencé bien après la publication par Neyman de la base théorique de ce type d'échantillonnage (1934). Les échantillons par quotas, qui doivent uniquement respecter des nombres cibles de personnes comportant des caractéristiques particulières, comme l'âge et le sexe, ont été utilisés pendant de nombreuses années, particulièrement pour des études de marché. Stephan et McCarthy (1958) examinent cette méthode d'échantillonnage non probabiliste dans les sondages électoraux et d'autres types d'enquêtes au milieu du XX^e siècle aux États-Unis.

Le type d'échantillonnage non probabiliste utilisé dans la recherche commerciale et les études de marché a changé de façon marquée au cours des 20 dernières années, l'accès à Internet s'étant répandu en Amérique du Nord et dans de nombreuses régions de l'Europe. Pendant les 10 dernières années, particulièrement, les enquêtes en ligne — dont les répondants sont tirés de panels de volontaires — sont devenues extrêmement populaires. La grande majorité de ces enquêtes ne reposent pas sur des échantillons probabilistes. La raison de leur popularité est le coût peu élevé par interview, de beaucoup inférieur même à celui des méthodes d'enquête sur échantillon probabiliste considérées comme à faible coût, tel l'envoi par la poste. Une partie de l'attrait des échantillons probabilistes s'est aussi perdue en raison de l'augmentation de la non-réponse (Brick et Williams, 2013), ainsi que des préoccupations concernant le sous-dénombrement de la base de sondage. Ces problèmes soulèvent des préoccupations en ce qui a trait à la validité des inférences à partir d'un échantillon probabiliste. Même les ardents défenseurs de l'échantillonnage probabiliste ont été forcés de se demander si un échantillon probabiliste comportant un faible taux de réponse ou de couverture conserve les propriétés très appréciées d'un échantillon probabiliste valide (Groves, 2006).

1. Westat et JPSM, 1600, Research Blvd, Rockville, Maryland, États-Unis 20850.

La section qui suit résume certaines constatations importantes d'un groupe de travail sur l'échantillonnage non probabiliste mis sur pied par l'American Association of Public Opinion Research (AAPOR). Elles servent de préambule à certaines méthodes actuelles et sujets de recherches futurs.

2. Rapport du groupe de travail

Le groupe de travail de l'AAPOR a été mandaté « pour examiner les cas où divers plans de sondage qui n'utilisent pas des échantillons probabilistes pourraient quand même être utiles pour faire des inférences en fonction d'une population plus importante ». Le rapport du groupe de travail, qui a été achevé au début de 2013, peut être téléchargé à partir du site Web de cette organisation (www.aapor.org). Baker et coll. (2013) ont résumé le rapport; les commentaires de cinq experts du domaine et une réplique sont publiés dans le même numéro de la publication. Plutôt que de reprendre à nouveau toutes les constatations, nous en avons choisi quelques-unes essentielles (entre guillemets ci-dessous), qui ont fait l'objet de plusieurs discussions par suite de la publication du rapport et de son sommaire.

« Contrairement à l'échantillonnage probabiliste, aucun cas unique n'englobe adéquatement tout l'échantillonnage non probabiliste. » L'argument de cet énoncé est parfois mal compris. L'objectif est de faire ressortir qu'il est peu valable de mettre toutes les méthodes non probabilistes dans le même panier, parce qu'elles diffèrent tellement. Les questions et les préoccupations concernant les méthodes d'échantillonnage axées sur le répondant et les panels de volontaires sur le Web sont très différentes. Même dans le domaine générique des panels de volontaires sur le Web, les méthodes utilisées pour sélectionner les répondants et produire des estimations peuvent différer.

« Les méthodes non probabilistes les plus prometteuses pour les enquêtes sont celles fondées sur des modèles visant à aborder les défis de l'inférence, à la fois aux étapes de l'échantillonnage et de l'estimation. » Cette constatation est plus hypothétique que fondée sur des résultats empiriques. Elle s'apparente de nombreuses façons à l'attente selon laquelle un plan de collecte dynamique peut entraîner un biais de non-réponse plus faible dans les échantillons probabilistes (Lundquist et Särndal, 2013). Cela vient du fait qu'un ensemble plus diversifié de répondants réduira les biais en présence d'un plan de pondération équivalent. Même si cela semble raisonnable, cela n'a pas encore été validé de façon uniforme dans des échantillons probabilistes (avec un plan de collecte dynamique) ou des échantillons non probabilistes.

« Si l'on veut que les échantillons non probabilistes soient mieux acceptés parmi les chercheurs du domaine des enquêtes, il doit y avoir un cadre plus cohérent et un ensemble de mesures connexes pour évaluer leur qualité. » Aucune étude ni série de travaux ne peut prouver qu'une stratégie de collecte de données et d'estimation produira des estimations qui sont raisonnables pour la plupart des usages. Par exemple, le *Literary Digest* avait prédit correctement le gagnant de chaque élection à partir de 1920, jusqu'à son erreur fatale de 1936, alors qu'il avait prédit l'élection de Landon. Les résultats empiriques sont importants, mais il doit exister un ensemble de principes qui appuie le processus de collecte des données et d'estimation, afin que les échecs puissent être expliqués. L'échantillonnage probabiliste repose sur un tel fondement, et la théorie fait en sorte que lorsque les estimations à partir d'un échantillon probabiliste ne sont pas exactes, les erreurs peuvent être liées à des écarts, comme la non-réponse, sans que la théorie doive être rejetée.

« Les échantillons non probabilistes pourraient convenir pour faire des inférences statistiques, mais la validité des inférences dépend du caractère approprié des hypothèses qui sous-tendent le modèle et de la façon dont les écarts par rapport à ces hypothèses influent sur les estimations particulières. » Selon les membres du groupe de travail, cette constatation est celle qui devrait susciter le plus de controverses (Bethlehem et Cooben, 2013). Même s'il s'agissait d'une question contentieuse au moment de la publication du rapport, nous avons pu déterminer que nombreux sont ceux qui étaient d'accord avec cette position, y compris la plupart des experts qui ont examiné l'article en question.

Il existe un autre domaine de la recherche statistique dont la situation s'apparente beaucoup à celle de l'échantillonnage non probabiliste, à savoir les études observationnelles. Selon Madigan et coll. (2014) :

« Les facteurs qui menacent la validité des études observationnelles sur les effets des interventions soulèvent des questions quant au rôle approprié de ce genre d'études dans la prise de décisions. Néanmoins, les revues scientifiques dans des domaines comme la médecine, l'éducation et les sciences sociales présentent bon nombre de ces études, souvent en explorant peu ces menaces, et la presse non spécialisée regorge de reportages basés sur ces études... le caractère introspectif et ponctuel de la conception de ces analyses semble passer outre toute évaluation objective significative de leur rendement... »

En dépit des préoccupations concernant la validité des études observationnelles, les chercheurs dans ce domaine comprennent l'importance essentielle du rôle de ces études et mettent l'accent sur la détermination de ce qui peut être fait pour améliorer la science. À notre avis, il faut le même sentiment d'urgence pour améliorer les échantillons non probabilistes, plutôt que d'en rejeter simplement toutes les formes comme n'étant pas fondées sur le plan scientifique.

Il existe des preuves que des travaux d'inférence à partir d'échantillons non probabilistes se poursuivent, même si la majeure partie d'entre eux sont plus empiriques que théoriques. Par exemple, Barratt, Ferris et Lenton (2014) utilisent un échantillon en ligne pour estimer la taille et les caractéristiques d'une sous-population rare. Leur méthode d'évaluation est similaire à celle de nombreuses études antérieures; ils comparent les estimations à partir d'un échantillon en ligne à celles reposant sur un échantillon probabiliste, et trouvent certaines différences importantes. Malgré ces différences, ils prétendent que l'échantillon en ligne peut être utile s'il est combiné à un échantillon probabiliste. Wang et coll. (2014) utilisent un échantillon d'utilisateurs de la console Xbox qui est clairement non représentatif des électeurs aux États-Unis et ont recours à des méthodes d'estimation fondées sur un modèle pour produire des prédictions électorales. Ils montrent que les estimations comportent de faibles biais, en dépit des problèmes que présente l'échantillon. Ces types d'applications et de recherches sont extrêmement utiles, même si aucune étude ne peut à elle seule fournir les fondements théoriques que nous considérons comme essentiels. Il est possible que des examens de nombreuses applications de ce genre contribuent à susciter de nouvelles façons de percevoir les questions de fondement.

3. Adaptation des données à leur utilisation et conditions d'utilisation

Le groupe de travail a également soulevé que l'utilité des estimations découlant d'échantillons non probabilistes (ou de tout échantillon par ailleurs) doit être liée directement à l'objectif visé, ce que l'on appelle « l'adaptation des données à leur utilisation ». Certaines façons d'évaluer si un échantillon non probabiliste devrait même être envisagé dans ce cas sont décrites ci-après, et elles sont liées à ce concept d'adaptation à l'utilisation. Nous nous attendons à ce que ces conditions changent au fur et à mesure que nous recueillerons plus de renseignements concernant des méthodes particulières d'échantillonnage non probabiliste et d'estimation.

La majeure partie des résultats empiriques a montré que les échantillons probabilistes comportent généralement des biais plus faibles que les échantillons non probabilistes (Callegaro et coll., 2014). Toutefois, il existe des cas où les échantillons non probabilistes peuvent représenter un meilleur choix. Selon Brick (2014), trois critères doivent être envisagés pour l'utilisation d'un échantillon non probabiliste plutôt que d'un échantillon probabiliste. Ces trois conditions sont les suivantes :

- a. le coût de la collecte des données pour l'échantillon non probabiliste doit être beaucoup plus faible que celui de l'échantillon probabiliste;
- b. les estimations ne doivent pas être extrêmement précises pour répondre aux exigences de l'enquête;
- c. lorsque la population cible est stable et bien comprise, un échantillon non probabiliste peut être envisagé, même lorsque des niveaux plus élevés d'exactitude sont nécessaires.

La condition a) est nécessaire, mais insuffisante. Autrement dit, il existe des cas où un échantillon non probabiliste à faible coût peut être pire que l'absence complète de données, parce qu'il donne lieu à des mesures contreproductives. Hansen, Madow et Tepping (1983) ont souligné que la différence de coût pour un échantillon fondé sur un modèle (échantillon non probabiliste) n'était pas beaucoup plus faible que pour un échantillon probabiliste, ce qui fait qu'il y avait peu de raisons de ne pas avoir recours à ce dernier. Internet a modifié la

structure des coûts de façon marquée depuis 1983; par conséquent, les coûts d'un échantillon non probabiliste peuvent maintenant être beaucoup plus faibles que ceux d'un échantillon probabiliste.

Les conditions b) et c) sont directement liées à l'adaptation des données à leur utilisation. Si l'on a besoin uniquement d'estimations approximatives d'une population, un échantillon non probabiliste à faible coût peut être approprié. Même dans les comparaisons qui ont fait ressortir que les estimations à partir d'un échantillon non probabiliste n'étaient pas aussi exactes que celles à partir d'échantillons probabilistes, les estimations à partir d'échantillons non probabilistes étaient similaires à celles d'échantillons probabilistes pour une vaste gamme de stratégies d'échantillonnage en ligne. La condition c) reconnaît que certains échantillons non probabilistes ont affiché une exactitude constante, mais il s'agit de cas où la population est stable et où il existe des données auxiliaires puissantes. Certaines enquêtes auprès d'établissements peuvent être classées dans cette catégorie, en raison de leur stabilité et de la présence de variables auxiliaires importantes dans la base d'échantillonnage (Knaub, 2007). Les études électorales aux États-Unis appartiennent aussi à cette catégorie, du fait particulièrement qu'elles sont des prédicteurs bien connus et puissants du comportement électoral. Il convient de souligner que le résultat d'une élection est unique et que les estimations pour d'autres caractéristiques à partir de ces échantillons non probabilistes n'ont pas été évaluées de près. Hansen (1987) émet une mise en garde : on ne peut pas présumer la stabilité lorsque la société ou la population cible connaît des changements.

La capacité de produire une vaste gamme d'estimations comportant des biais faibles ou raisonnables figure parmi les caractéristiques qui font que les résultats empiriques d'un échantillon probabiliste se démarquent de façon frappante de ceux des échantillons non probabilistes en général. C'est cette capacité polyvalente qui manque le plus aux échantillons non probabilistes. Cela s'explique entre autres par l'activité de modélisation requise dans les échantillons non probabilistes. Par exemple, il est pratique de modéliser soigneusement un résultat particulier (dans les études électorales, il pourrait s'agir de plusieurs contestations d'élections qui comportent des rapports similaires entre les prédicteurs et le résultat), et de le faire avec plus de précision et peut-être moins de biais qu'au moyen d'un échantillon probabiliste type. On a recours au même type d'effort de modélisation dans les modèles d'estimation sur petits domaines, lorsque la taille de l'échantillon est trop petite, et parfois même nulle, pour produire des estimations fiables à partir de méthodes fondées sur un plan de sondage. Nous ne connaissons pas de modèles d'estimation sur petits domaines qui ont été proposés pour une vaste gamme de statistiques et de fins. Cela présente un défi important pour les échantillons non probabilistes reposant sur des échantillons en ligne.

Ces problèmes, qui sont inhérents aux échantillons non probabilistes, ne signifient pas que l'échantillonnage probabiliste n'a pas ses propres lacunes graves. Les hypothèses de l'échantillonnage probabiliste ne se vérifient pas en pratique, et les erreurs de couverture, ainsi que les erreurs dues à la non-réponse, sont au cœur de cet échec. Les erreurs de mesure, qui sont souvent même plus importantes, touchent à la fois les échantillons probabilistes et non probabilistes. Par exemple, les estimations empiriques des différences par rapport aux totaux de population figurant dans Messer et Dillman (2011) montrent clairement qu'un échantillon probabiliste ne peut pas à lui seul immuniser les estimations d'enquête contre les biais importants.

4. Méthodes d'échantillonnage non probabiliste en ligne

Le groupe de travail examine et définit une gamme variée de méthodes d'échantillonnage en ligne utilisées dans les échantillons non probabilistes. Callegaro et coll. (2014) les abordent de façon plus détaillée, ce que nous ne ferons pas ici. De nombreuses enquêtes volontaires utilisent des méthodes d'échantillonnage de type « river » et « router », qui permettent de tirer des répondants de diverses pages Web et de les intégrer à une enquête. Par conséquent, ces méthodes et des méthodes similaires sont sujettes à des biais de sélection qui sont très complexes, et les tentatives en vue de compenser les biais, comme le recours aux ajustements par pondération par la propension, laissent à désirer.

Les panels Web qui utilisent des répondants obtenus grâce à ces types de méthodes d'échantillonnage sont sujets à des biais provenant exactement des mêmes sources. Les panels comportent des caractéristiques avantageuses additionnelles. Par exemple, les données du « profil » des membres du panel peuvent servir à rajuster les estimations, et cela peut être utile pour traiter la non-réponse d'un panel. Cela pourrait mener à des biais plus faibles dans les estimations de changement. Ces « profils » de panel peuvent aussi être utiles pour réduire les biais de

sélection originaux, mais cela reste à prouver dans une large mesure. Les biais de sélection sont très difficiles à comprendre et à résoudre efficacement.

Dans la foulée des travaux d'études observationnelles et épidémiologiques, on a commencé à utiliser l'appariement dans certaines enquêtes volontaires (Rivers et Bailey, 2009). Comme il a été mentionné précédemment, le groupe de travail de l'AAPOR considérait cette approche comme très prometteuse, mais les résultats peu reluisants des études observationnelles doivent être pris en compte. La capacité d'utiliser l'appariement en combinaison avec d'autres méthodes d'échantillonnage et de pondération en ligne présente un défi, en raison de la nécessité d'estimer un si grand nombre de caractéristiques et de relations dans les enquêtes. Dans les études observationnelles, il y a généralement un nombre beaucoup plus faible de statistiques relatives aux résultats clés, et la faisabilité d'une modélisation soigneuse de chaque résultat est grandement améliorée.

Une autre approche en ce qui a trait aux méthodes d'échantillonnage non probabiliste a été explorée, à savoir combiner un échantillon non probabiliste important en ligne avec un petit échantillon probabiliste (souvent appelé échantillon de référence), afin de procéder à des rajustements pour tenir compte des biais dans l'échantillon non probabiliste. Le principal obstacle que pose cette méthode est que la taille réelle de l'échantillon est fonction de l'échantillon probabiliste, ce qui fait que le grand échantillon non probabiliste est essentiellement réduit à la taille de l'échantillon probabiliste. Ces questions sont abordées par plusieurs chercheurs, et remontent à plusieurs années, au moins jusqu'aux travaux de Cochran (1977) et, plus récemment, ceux de Bethlehem (2008). Dever, Rafferty et Valliant (2008) montrent que cette approche peut réduire les biais, mais renforce la perte de précision associée à l'échantillon de référence. Il existe une approche connexe, soit une combinaison hybride d'échantillons probabilistes et non probabilistes (Berzofsky, Williams et Biemer, 2009); toutefois, cette méthode n'a pas suscité beaucoup d'intérêt jusqu'à maintenant.

5. Voie à suivre

L'avenir de l'échantillonnage en ligne n'est pas clair, et les multiples approches qui ont été tentées au cours des 10 dernières années témoignent de sa nature en évolution. Durant cette période, on a aussi effectué des recherches sur l'échantillonnage probabiliste, afin de résoudre les problèmes qu'il comporte. Parmi les voies à suivre pour l'avenir, on pourrait tirer parti de tous ces efforts pour améliorer le rendement empirique et la base théorique des deux méthodologies.

Par exemple, la recherche sur l'échantillonnage et la collecte des données, comme l'échantillonnage équilibré, le plan de collecte dynamique, le plan de sondage adaptatif, les indicateurs R et d'autres mesures de la représentativité, pourrait être une façon de trouver de meilleures méthodes pour les deux types d'études. Les ajustements de la couverture, les totaux de contrôle estimés et les ajustements non ignorables pour tenir compte de la non-réponse pourraient être des méthodes d'estimation comportant une importance pratique.

Outre ces méthodes d'enquête, d'autres domaines offrent des idées qui pourraient être envisagées. L'inférence causale a fait ses preuves pour résoudre le biais de sélection, et des méthodes plus nouvelles sont sans cesse mises en œuvre et explorées. Des domaines comme la psychologie cognitive et la recherche comportementale ont aussi gagné en importance par rapport au moment où ils ont été adoptés pour la première fois comme outils pour la recherche d'enquête dans les années 1970. Évidemment, la science de l'information a connu une révolution, et de nouveaux domaines, comme les mégadonnées, et des anciens concepts assortis de nouvelles technologies, comme les dossiers administratifs, pourraient apporter de nouvelles idées et doivent être envisagés.

Même s'il ne s'agit pas d'une recette pour améliorer l'inférence basée sur l'échantillonnage non probabiliste, cela montre que la recherche est possible et essentielle. Il existe des outils et des méthodes qui pourraient fournir le cadre des inférences à partir d'échantillons non probabilistes, mais sans travaux de recherche innovateurs, la confusion actuelle subsistera.

Bibliographie

- BAKER, Reg, J. Michael BRICK, Nancy A. BATES, Mike BATTAGLIA, Mick P. COUPER, Jill A. DEVER, Krista J. GILE et Roger TOURANGEAU. 2013. « Summary Report of the AAPOR Task Force on Non-probability Sampling », *Journal of Survey Statistics and Methodology*, vol. 1, p. 90 à 143.
- BARRATT, Monica J., Jason A. FERRIS et Simon LENTON. 2014. « Hidden Populations, Online Purposive Sampling et External Validity Taking off the Blindfold », *Field Methods*, 1525822X14526838.
- BERZOFSKY, Marcus, Rick WILLIAMS et Paul BIEMER. 2009. « Combining Probability and Non-Probability Sampling Methods: Model-Aided Sampling and the O*NET Data Collection Program », *Survey Practice* 2.6. Téléchargé le 15 décembre 2015 à partir de www.surveypactice.org/index.php/SurveyPractice/article/view/184/html.
- BETHLEHEM, Jelke. 2008. « Peut-on établir des statistiques officielles à partir d'enquêtes en ligne reposant sur le principe de l'autosélection? », *Recueil du Symposium de Statistique Canada*.
- BETHLEHEM, Jelke, et Fannie COOBEN. 2013. « Web Panels for Official Statistics? », *Proceedings 59th ISI World Statistics Congress, 25 au 30 août 2013, Hong Kong*. Téléchargé le 1^{er} mai 2014 à partir de <http://2013.isiproceedings.org/Files/IPS064-P1-S.pdf>.
- BRICK, J. Michael. 2014. « On Making Inferences from Non-Probability Samples ». Washington Statistical Society 2014 President's Invited Seminar, Washington DC (26 mars 2014).
- BRICK, J. Michael, et Douglas Williams. 2013. « Explaining rising nonresponse rates in cross-sectional surveys », *ANNALS of the American Academy of Political and Social Science*, vol. 645, p. 36 à 59.
- CALLEGARO, Mario, R. BAKER, J. BETHLEHEM, A. GÖRITZ, J. KROSNICK et P. LAVRAKAS, publié sous la direction de. 2014. *Online Panel Research: A Data Quality Perspective*, John Wiley & Sons.
- COCHRAN, William. 1977. *Sampling Techniques*, John Wiley & Sons, New York.
- DEVER, Jill A., Ann RAFFERTY et Richard VALLIANT. 2008. « Internet Surveys: Can Statistical Adjustments Eliminate Coverage Bias? », *Survey Research Methods*, vol. 2, p. 47 à 62.
- GROVES, Robert M. 2006. « Nonresponse rates and nonresponse bias in household surveys », *Public Opinion Quarterly*, vol. 70, p. 646 à 675.
- HANSEN, Morris H. 1987. « Some History and Reminiscences on Survey Sampling », *Statistical Science*, vol. 2, n° 2, p. 180 à 190.
- HANSEN, Morris H., William N. HURWITZ et William G. MADOW. 1953. *Sampling Survey Methods and Theory. Vol I*, John Wiley & Sons, New York.
- HANSEN, Morris H., William G. MADOW et Benjamin J. TEPPING. 1983. « An evaluation of model-dependent and probability-sampling inferences in sample surveys », *Journal of the American Statistical Association*, vol. 78, n° 384, p. 776 à 793.
- KNAUB James R. 2007. « Cutoff Sampling and Inference », *InterStat*, avril.
- LOHR, Sharon. 2009. *Sampling: Design and Analysis*, Cengage Learning.
- LUNDQUIST, Peter, et Carl-Erik SÄRNDAL. 2013. « Aspects of responsive design with applications to the Swedish Living Conditions Survey », *Journal of Official Statistics*, vol. 29, p. 557 à 582.

- MADIGAN, David, P. STANG, J. BERLIN, M. SCHUEMIEJ, M. OVERHAGE, M. SUCHAR, B. DUMOUCHEL, A. HARTZEMA, P. RYAN. 2014. « A Systematic Statistical Approach to Evaluating Evidence from Observational Studies », *Annual Review of Statistics and Its Application*, vol. 1, p. 11 à 39.
- MESSER, Benjamin L., et Don A. DILLMAN. 2011. « Surveying the general public over the Internet using address-based sampling and mail contact procedures », *Public Opinion Quarterly*, vol. 75, p. 429 à 457.
- NEYMAN, Jerzy. 1934. « On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection », *Journal of the Royal Statistical Society*, vol. 97, p. 558 à 625.
- RIVERS, Douglas, et Delia Bailey. 2009. « Inference from Matched Samples in the 2008 U.S. National Elections ». Document présenté à la 64^e conférence annuelle de l'American Association for Public Opinion Research, Hollywood, Floride, mai.
- SÄRNDAL, Carl-Erik, Bengt SWENSSON et Jan WRETMAN. 1992. *Model Assisted Survey Sampling*, Springer.
- SQUIRE, Peverill. 1988. « Why the 1936 Literary Digest Poll Failed? », *Public Opinion Quarterly*, vol. 52, p. 125 à 133.
- STEPHAN, Fredrick F., et Philip J. MCCARTHY. 1958. *Sampling Opinions: An Analysis of Survey Procedure*, John Wiley & Sons, New York.
- WANG, Wei, David ROTHSCHILD, Sharad GOEL et Andrew GELMAN. 2014. « Forecasting Elections with Non-Representative Polls », *International Journal of Forecasting*, doi:10.1016/j.ijforecast.2014.06.001.