

Article

Dans ce numéro



Juin 1999

Dans ce numéro

Chers lecteurs,
Chères lectrices,

Je m'empresse de partager avec vous deux bonnes nouvelles. Tout d'abord, le numéro de décembre marquera le 25^e anniversaire de *Techniques d'enquête*. Ce sera un numéro un peu plus volumineux de la revue, comportant des articles d'éminents statisticiens contemporains. Deuxièmement, nous comptons préparer une version électronique de la revue. Nous espérons monter le numéro de décembre 1999 sur un site Web spécial. Tous les abonnés actuels pourront télécharger la revue gratuitement. Les résultats de cet essai nous permettront de déterminer s'il serait préférable d'offrir la version électronique au lieu de la version imprimée, ou en plus de celle-ci. Le prochain numéro vous fournira des renseignements supplémentaires. Vos remarques et vos suggestions sont toujours bien accueillies.

Le présent numéro aborde un choix de sujets : trois articles sur les estimations infra-régionales, quatre sur les estimations en général, deux sur les nouveaux plans d'échantillonnage et deux autres sur l'analyse des données.

Kröger, Särndal et Teikari proposent une nouvelle famille de plans d'échantillonnage, sous la manchette d'échantillonnage mixte de Poisson, qui correspond à un mélange pondéré de plans d'échantillonnage de Poisson et de Bernoulli. Grâce à une étude de Monte Carlo portant sur des données finlandaises, ils montrent empiriquement que, pour divers estimateurs ponctuels, l'échantillonnage mixte de Poisson est plus efficace que l'échantillonnage de Poisson ordinaire.

Bell et Kramer cherchent à résoudre le problème que pose depuis longtemps l'estimation de la variance des estimateurs X-11. Chaque mois, les bureaux de la statistique du monde entier publient les estimations brutes de diverses variables, ainsi qu'une mesure correspondante de l'erreur, ordinairement l'erreur-type ou le coefficient de variation. Par contre, ils ne sont en mesure de produire aucune mesure de ce genre pour les estimations désaisonnalisées ou estimations de tendance correspondantes obtenues par la méthode X-11. Bell et Kramer proposent une méthode intéressante qui permet de résoudre pratiquement le problème. Ils calculent deux formes d'erreur, l'une résultant de l'erreur d'échantillonnage et l'autre, de l'utilisation d'extrapolations ARMMI pour prolonger la série aux deux extrémités.

De Haan, Opperdoes et Schut se penchent sur l'échantillonnage d'articles dans un groupe de produits, aux fins de leur inclusion dans le calcul de l'indice des prix à la consommation, d'après les données de lecteurs optiques. À l'heure actuelle, la plupart des bureaux de la statistique appliquent une méthode d'échantillonnage raisonné qui biaise naturellement les estimations. Les auteurs s'efforcent de déterminer si l'échantillonnage probabiliste serait une meilleure méthode en ce qui concerne l'erreur quadratique moyenne et obtiennent des résultats intéressants.

Pierre Duchesne considère une nouvelle classe d'estimateurs de calage robustes qui permet d'obtenir des poids restreints à un intervalle donné. La démarche consiste à transformer en poids cales des poids implicites robustes judicieusement choisis. Dans une brève étude empirique, les nouveaux estimateurs sont à la fois illustrés et comparés avec des estimateurs déjà proposés.

Tillé étudie une méthode d'échantillonnage répété qui tient compte de renseignements auxiliaires. Il commence par généraliser l'utilisation des probabilités conditionnelles d'inclusion de façon à pouvoir les appliquer à tout plan d'échantillonnage. Puis, il construit des estimateurs que l'on peut considérer comme étant des estimateurs linéaires optimaux et les compare à l'estimateur de régression généralisé (GREG). Il différencie les divers estimateurs grâce à un ensemble de simulations. Enfin, il discute du problème de l'interaction entre le plan de sondage et les variables auxiliaires.

Prasad et Rao étudient le problème de l'estimation régionale au moyen d'un modèle à effets aléatoires. Alors que les méthodes classiques d'estimation des moyennes régionales sont fondées sur un modèle, Prasad et Rao obtiennent des estimations fondées sur le plan de sondage (assistées par modèle) grâce à l'intégration des poids d'enquête. Ils calculent aussi les estimateurs à base de modèle correspondants de l'erreur quadratique moyenne (EQM) des estimations régionales. Ils montrent, grâce à une simulation, que leur estimateur de l'EQM est assez stable et n'est que légèrement biaisé.

Dans leur article sur l'estimation régionale, Moura et Holt se concentrent sur des modèles à plusieurs niveaux, qui tiennent compte de données auxiliaires au niveau tant de l'unité que de la petite région, ainsi que d'effets aléatoires régionaux en ce qui concerne aussi bien les coordonnées à l'origine que la pente des droites de régression. Ils estiment les paramètres des modèles à effets fixes et aléatoires par une méthode itérative généralisée des moindres carrés restreinte. Ils calculent une valeur approchée de l'erreur quadratique moyenne. Ils montrent, par des simulations, que le modèle permet de produire de meilleurs estimateurs régionaux que ceux fondés sur des modèles plus simples, qu'une spécification trop poussée du modèle n'entraîne aucune perte importante d'efficacité et que l'approximation de l'EQM et l'estimateur connexe de l'EQM donnent de bons résultats.

Chattopadhyay, Lahiri, Larsen et Reimnitz étudient l'estimation des proportions régionales d'événements rares. Ils se servent des données d'une enquête téléphonique sur la consommation d'alcool et de drogues pour illustrer leur méthode et la comparer à d'autres. L'estimateur qu'ils proposent combine des estimations démographiques par groupe âge/sexe/comté fondées sur des données de recensement et des estimations bayésiennes empiriques des proportions dans ces groupes fondées sur des données d'enquête. Ils proposent un estimateur jackknife de l'erreur quadratique moyenne qui tient compte de la variabilité due à l'estimation des paramètres du modèle.

Rubin-Bleuer et Kovačević étudient les problèmes que pose l'estimation des proportions longitudinales de faibles revenus d'après les données d'une enquête longitudinale dont le plan de sondage est complexe. Ils envisagent deux estimateurs fondés sur le plan de sondage : l'un s'appuyant à la fois sur les échantillons longitudinal et transversal, appelé « estimateur mixte » et l'autre, uniquement sur l'échantillon longitudinal. Grâce à une simulation, ils comparent les deux estimateurs en cas d'érosion de l'échantillon en appliquant des modèles de correction pour l'érosion fondés sur l'hypothèse, d'une part, que les unités d'échantillonnage « manquent au hasard » et, d'autre part, qu'elles « manquent complètement au hasard ». Ils illustrent les résultats au moyen de données de deux enquêtes longitudinales.

Tate étudie le couplage des données recueillies sur une même personne au cours de cycles trimestriels subséquents de la British Labour Force Survey, une enquête par panel avec renouvellement d'un cinquième de l'échantillon à chaque cycle. Elle examine les divers facteurs susceptibles de biaiser les analyses réalisées d'après des données couplées de ce genre. Plus précisément, elle étudie les effets éventuels de l'érosion de l'échantillon, des erreurs de réponse et de la réponse par procuration. Elle envisage aussi diverses méthodes de correction pour tenir compte de ces biais.

Enfin, dans une note brève, Gabler, Haeder et Lahiri présentent une justification fondée sur un modèle de la formule bien connue de l'effet de plan de sondage proposée par Kish. Ils montrent que la formule produit en réalité une valeur prudente de l'effet réel de plan de sondage.

Le rédacteur en chef