
Échantillonnage : théorie, méthodes et applications

Président: Changbao Wu (University of Waterloo)

MIN CHEN, University of Waterloo

Méthode de pseudo-vraisemblance empirique pour les études prétest/post-test avec données d'enquête

Les études prétest/post-test représentent une méthode importante et populaire pour évaluer l'effet d'un traitement. La méthode de vraisemblance empirique est un outil d'inférence utile pour de telles études dans lesquelles des informations sur un ensemble de variables initiales peuvent être intégrées de façon efficace. Nous étendons l'application de la méthodologie aux études sociales ou psychologiques avec des données recueillies par le biais d'enquêtes longitudinales complexes. Notre méthode basée sur la pseudo-vraisemblance empirique peut être utilisée pour évaluer le changement de comportement des gens ou les schémas de réponse découlant de la mise en œuvre de nouvelles politiques publiques ou d'un changement dans le protocole de collecte des données. La méthode est illustrée à l'aide des données du Projet international d'évaluation de la lutte antitabac.

GAURI SANKAR DATTA, University of Georgia, U.S. Bureau of the Census

Approche de vraisemblance pondérée pour l'estimation de modèles pour petits domaines avec données au niveau des unités

L'estimation pour petits domaines fait appel à des modèles au niveau des régions ou des unités. Les modèles au niveau des régions s'appliquent aux estimés d'enquêtes directs généralement convergents sous le plan, donnant ainsi des prévisions du modèle convergents également sous le plan. Cependant, les modèles au niveau des unités s'appliquent en général aux microdonnées d'enquête en ignorant les poids d'échantillonnage, ne conduisant pas à des prévisions convergentes sous le plan. Kott, Rao et d'autres ont développé des méthodes pour intégrer les poids d'échantillonnage dans le modèle de régression avec erreurs normales imbriquées au niveau de l'unité afin d'atteindre la convergence sous le plan. Nous proposons une pseudo-vraisemblance incorporant les poids d'échantillonnage qui s'applique à des données normales ou non normales, y compris les données binaires et de dénombrement. En utilisant une approche bayésienne, nous appliquons notre méthode aux données de l'American Community Survey.

DARREN GRAY, Statistics Canada

Une répartition d'échantillon sous contraintes multiples pour des estimations nationales détaillées complexifiée par l'ajout de demandes de partenaires provinciaux

Pour l'enquête sur l'innovation et les stratégies d'entreprise, l'allocation de l'échantillon en 2009 devait satisfaire les exigences de précision pour trois modules de questions ciblant chacun des combinaisons d'industries et groupes de taille d'emploi différents au niveau national. Pour l'enquête de 2012, ce système de contraintes déjà complexe s'est complexifié davantage par l'ajout de demandes provinciales. Une méthode de programmation non-linéaire a été proposée pour résoudre ce problème. Cet article discutera des méthodes employées pour s'assurer que toutes les exigences soient satisfaites tout en minimisant la taille d'échantillon et en évitant les strates de petite taille.

PIERRE-OLIVIER JULIEN, Statistique Canada

Coordination négative d'échantillons: un survol de quelques méthodes

Une quantité considérable de données est produite par les agences statistiques. Ceci ajoute une charge de travail parfois importante aux répondants. Plusieurs méthodes de coordination d'échantillons tirés à même une base de sondage unique furent développées pour réduire leur fardeau de réponse. On échafaude présentement à Statistique Canada un projet central qui couvrira plus de 100 enquêtes basées sur le Registre des entreprises. La coordination de tous ces échantillons pourrait être possible à partir d'un outil informatique commun. Cette présentation se veut un survol de certaines méthodes qui pourraient être implantées dans un tel outil.

CHRISTIAN OLIVIER NAMBEU, Statistique Canada

Calage utilisant des contraintes estimées

Le calage est une approche d'estimation intensément utilisée en pratique à cause de ses nombreux avantages. Très souvent, les totaux de calage sont des quantités connues au niveau de la population. Dans cet article, nous étudions le cas où certains totaux sont remplacés par des totaux estimés. L'estimateur correspondant est analysé et comparé à des estimateurs par calage habituels en termes de biais et d'erreur quadratique moyenne théoriquement et via une étude par simulations.