

ESTIMATION DES TRAFICS VOYAGEURS SNCF TRANSILIEN : UN PLAN DE SONDAGE COMPLEXE ET PRISE EN COMPTE D'ÉCHANTILLONS MULTIPLES

Anaïs Rémy ¹ & Maguelonne Chandesris ¹

¹ *SNCF Innovation & Recherche, 40 avenue des Terroirs de France, 75611 PARIS,*
anaïs.remy@sncf.fr, maguelonne.chandesris@sncf.fr

Connaître précisément les flux de voyageurs fait partie des fondamentaux d'un opérateur de transporteur. Cela lui permet de proposer une offre adaptée au mieux aux comportements et attentes de ses clients (suivi et organisation de la production, définition des plans d'investissement adéquats ...). En Île-de-France, la multiplicité des acteurs présents sur le marché et l'utilisation de titres de transports forfaitaires (mensuel, annuel) rend difficile l'évaluation fine du volume de voyageurs transportés par la SNCF. Dans ce contexte, SNCF Transilien désire disposer d'un nouvel outil d'analyse des trafics voyageurs sur différents agrégats (global, par ligne, mensuel, journalier, etc.).

Depuis plusieurs années, de nombreuses technologies (contrôles automatiques de billets (CAB), comptages automatiques de voyageurs, ventes de titres de transport ...) mettent à la disposition de la SNCF de véritables gisements de données pour l'évaluation des trafics voyageurs. Pour autant, ces données sont actuellement complexes à analyser au vu notamment des volumétries importantes à manipuler (de l'ordre de 9 millions de gestes de validations enregistrés en Ile-de-France par jour) et de la connaissance parcellaire de l'information liée à un équipement non exhaustif : par exemple l'ensemble des gares Transilien ne sont pas équipées en contrôles automatiques de billets.

L'objectif de ce travail est de développer des méthodes d'estimations des trafics voyageurs à partir de l'exploitation de données recueillies par divers systèmes (ventes, comptages manuels et automatiques, validation issues des CAB, enquêtes ...). La construction des estimateurs soulève deux problématiques :

- Les données historiques récupérées peuvent être vues comme des échantillons de données déjà tirés selon un plan de sondage à définir. Pour cela, nous avons proposé une modélisation a posteriori du plan de sondage réalisé (plan stratifié à plusieurs degrés).
- D'autre part, les méthodes de mise en commun de données multiples sont utilisées afin de tirer partie de la richesse des données disponibles, et ainsi consolider les résultats.

Bibliographie

Ardilly, P. (2006). Les techniques de sondage, Edition Technip.

Deville, J-C., et Maumy, M. (2006). Extension de la méthode d'échantillonnage et son application aux enquêtes dans le tourisme, Techniques d'enquête, 32, 197-206.

Lavallée, P. (2007). Indirect Sampling, Springer series in statistics.

Le Guennec, J. (2009). Les extensions régionales et locales de l'enquête logement 2006 - Echantillonnage et repondération, Actes JMS 2009.