

SÉMINAIRE INED DE STATISTIQUE APPLIQUÉEⁱ

ANALYSE DE TRAJECTOIRES

Séance du mardi 8 avril 2008 14h - 17h 30

L'analyse de trajectoires permet d'étudier l'évolution d'un événement au sein d'une population de référence. Les données utilisées peuvent avoir été recensées de manière prospective (techniques de panels), ou de manière rétrospective (enquêtes de type biographique). L'individu se déplace alors au cours du cycle d'étude dans une configuration où il oscille entre plusieurs états exclusifs. L'évolution de sa trajectoire, de son parcours, peut alors être représentée comme une succession de couples (E_i, T_i) , E_i représentant un état et T_i une date de changement d'état.

L'étude descriptive peut alors être effectuée par des méthodes relevant plutôt de l'analyse de données : analyses factorielles, classifications, appariement optimal. Par exemple le parcours professionnel ou la trajectoire résidentielle de l'individu peut être étudié, et les trajectoires individuelles regroupées selon leur similarité. Ainsi, les typologies de trajectoires permettront de définir des parcours-types.

Ce sont ces méthodes exploratoires que les différents intervenants vont expliciter et indiquer leur mise en œuvre au cours de cette demi-journée.

RÉSUMÉS DES INTERVENTIONS

Yvette GRELET Yvette, CEREQ

Les typologies de parcours, méthodes et usages •

La présentation brossera un panorama des méthodes de classification de trajectoires, et traitera des questions relatives au codage des données, à la représentation et à l'interprétation des classes, à partir de l'exemple des calendriers mensuels d'emploi des enquêtes Génération du Céreq.

Patrick ROUSSET, CEREQ

Typologie des parcours d'insertion à partir des cartes d'auto-organisation

Les calendriers professionnels dans les enquêtes longitudinales fournissent une information détaillée sur l'évolution de la situation professionnelle des individus en permettant notamment de recueillir la durée passée dans chaque situation et les transitions entre ces situations. De nombreux travaux ont proposé des méthodes statistiques pour réduire cette information en essayant de regrouper et de classer l'ensemble des parcours individuels en un certain nombre de trajectoires représentatives. Nous proposons dans cette communication le calcul d'une distance entre les parcours qui permet de tenir compte de la proximité des certaines situations professionnelles au cours du temps. Nous utilisons ensuite pour la classification des cartes d'auto-organisation qui permet de travailler sur un nombre de classes plus important que les méthodes classiques. Une application est proposée à partir des

données d'une enquête du CEREQ, «Génération 98» permettant de suivre le parcours professionnel des jeunes sur leurs sept premières années de vie active.

France GUÉRIN-PACE, INED •

Etude des parcours professionnels par les méthodes de la statistique textuelle.

Les méthodes de la statistique textuelle sont nées de la rencontre entre plusieurs disciplines : la linguistique, l'analyse du discours, la statistique, l'informatique, le traitement des enquêtes (L. Lebart, A. Salem, 1994). Elles sont utilisées pour traiter des corpus de textes dans des domaines très variés : linguistique, histoire, marketing, psychologie, sociologie... Une première famille d'applications privilégie l'analyse du contenu des textes sur leur forme. Elle s'applique aux traitements de questions ouvertes dans les enquêtes, entretiens, interviews, discours. Un second type d'application de ces méthodes s'intéresse à la comparaison de textes sur la base d'une étude quantitative du vocabulaire. Les applications sont nombreuses : attribution d'écrits historiques ou littéraires, étude ou comparaison du style de plusieurs auteurs, évolution du style, etc. Ces méthodes s'avèrent aussi d'une grande utilité pour l'étude d'itinéraires recueillis dans les enquêtes. On considère alors une séquence d'états (emploi, statut professionnel, lieux de résidence..) comme une suite de mots artificiels qui constituent une phrase. Comparé à des corpus littéraires, le vocabulaire du corpus est nécessairement pauvre mais il correspond à une diversité de situations importante. L'intérêt majeur de l'utilisation de ces méthodes à ce type de corpus réside dans la possibilité de comparer et d'analyser des parcours de longueur différente dans lesquels les personnes n'ont pas effectué le même nombre d'étapes.

Nicolas ROBETTE, INED

Optimal matching : théorie et application aux trajectoires professionnelles de franciliens.

L'étude des parcours de vie dans les sciences sociales fait l'objet depuis plusieurs décennies de très nombreux travaux. D'un point de vue méthodologique, ce sont principalement les approches probabilistes, centrées sur la modélisation d'évènements (approche atomiste), qui ont été privilégiées. Toutefois, la complexité croissante des parcours a amené le développement de techniques exploratoires, appréhendant la trajectoire comme un tout (approche holiste). Si diverses méthodes ont été proposées, en particulier dans le champ de l'analyse de séquences, c'est l'Optimal Matching Analysis qui bénéficie du plus large écho. Issue de la biologie moléculaire puis introduite dans les sciences sociales par Andrew Abbott dans les années 1980, l'OMA consiste à mesurer la similarité de chaque paire de séquences puis à construire à partir de cette mesure une matrice de distance, utilisable par les méthodes de classification.

Arnaud Bringé et Bénédicte Garnier

ⁱ Institut national d'études démographiques • Établissement public scientifique et technologique

133, bd Davout, Paris 20^e • Standard : 01 56 06 20 00 • Fax : 01 56 06 21 99

Le service Méthodes Statistiques de l'Ined organise un séminaire trimestriel, se proposant de couvrir un domaine de Statistique Appliquée traitant de l'application d'une méthodologie.

Ce séminaire est ouvert à tous, sans frais de participation. Pour tout autre renseignement contacter Madame Chafika Mekhazni, secrétaire du service, au 01 56 06 20 91.