

SÉMINAIRE INED LES RENCONTRES DE STATISTIQUE APPLIQUÉE

ANALYSE BIOGRAPHIQUE, ENTRE THÉORIE ET PRATIQUE

Séance du 28 avril 2009 de 14h à 17h30

L'approche biographique s'applique à des données longitudinales. Elle permet notamment de modéliser la survenue d'un événement démographique en fonction de caractéristiques individuelles. La mise en œuvre de cette étape de modélisation sera discutée au cours de cette séance à travers plusieurs réflexions.

RÉSUMÉS DES INTERVENTIONS

14h15 – Analyse biographique, du questionnaire à l'analyse – Éva LELIÈVRE • INED

Présentation d'introduction à la séance.

14h40 – Modèle d'analyse des biographies en temps discret – Jean-Marie LEGOFF

• Université de Lausanne

Les modèles les plus courants de l'analyse des biographies, notamment le modèle de Cox, reposent sur l'idée que les transitions analysées sont mesurées sur un temps continu. Dans la pratique, le temps est toujours mesuré sur une échelle discrète et les modèles en temps continu sont utilisés lorsque l'unité de mesure du temps est petite, par exemple le mois dans le cas des événements sociodémographiques. Dans certains cas, toutefois, l'unité de temps est plus longue, par exemple l'année, et il devient alors plutôt intéressant d'estimer des modèles appartenant à la famille des modèles en temps discret. Par rapport aux modèles à temps continu, ce n'est plus le taux instantané (*hazard rate*) qui fait l'objet d'une modélisation, mais la probabilité conditionnelle de connaître un événement au cours d'une unité de temps. Dans cette communication, nous présenterons deux modèles à temps discret qui sont les modèles logistique et complémentaire *loglog* en temps discret. Après une brève présentation théorique de ces modèles, seront discutées les modalités pratiques de leur application, telles que la préparation des données, leur estimation et la lecture des résultats.

15h45 – Analyse exploratoire d'un modèle explicatif – Application au modèle de Cox –

Xavier BRY • Université de Montpellier, Philippe ANTOINE • IRD-CEPED, chercheur associé à l'Ined.

Dans la modélisation explicative, le choix des prédicteurs est une étape complexe et délicate. Les contraintes, partiellement contradictoires, qui conditionnent cette étape sont: 1) la nécessité de tenir compte d'un modèle conceptuel qui traduise le point de vue de l'analyste et incorpore sa structuration logique du phénomène; 2) la nécessité de prendre en compte la possibilité de dimensions prédictives non prévues a priori; 3) l'éloignement des multicolinéarités; 4) la richesse du modèle, i.e. la nécessité de laisser intervenir tous les prédicteurs non étrangers au phénomène; 5) la parcimonie du modèle; 6) la robustesse du modèle.

L'Analyse en Composantes Thématiques est une méthode factorielle qui tente de satisfaire ces contraintes. Après en avoir exposé le principe dans le cadre linéaire classique, nous en proposons une extension à la modélisation linéaire généralisée, en particulier au modèle de Cox. La méthode est appliquée à l'étude du risque de divorce des hommes à Dakar. D'autres possibilités d'applications seront évoquées et discutées.

16h25 – Modèle de durée et prise en compte de l'hétérogénéité non observée – Théorie et application au marché du travail des jeunes.

Vanessa DI PAOLA •, Université de la Méditerranée (Laboratoire d'Économie et de Sociologie du Travail), Éric CAHUZAC • INRA Toulouse

La présentation propose de faire le point sur la question de la présence de l'hétérogénéité non observée dans les modèles de durée.

En effet, les modèles de durée autorisent l'évaluation du temps passé dans un état, en prenant en compte la présence de données censurées et le fait que l'individu était encore dans l'état à l'instant d'avant. Ce dernier point revient à considérer le phénomène de dépendance temporelle. Ainsi, il est possible à travers ces modèles, d'évaluer le rôle des caractéristiques individuelles et celui du temps même passé dans l'état, sur les chances d'en sortir (chômage, emploi, situation matrimoniale...). Pour autant, l'identification de la part "réelle" des caractéristiques individuelles et de celle de la dépendance temporelle ne va pas sans poser problème à cause de la présence incontournable d'hétérogénéité non observée, c'est-à-dire de caractéristiques individuelles influençant le phénomène considéré mais non prise en compte dans le modèle. Il convient donc de contrôler cette hétérogénéité non observée afin que les estimateurs obtenus ne soient pas biaisés (phénomène /mover-stayer/). C'est ce que nous développerons à travers des exemples pris sur le marché du travail.

Arnaud Bringé et Bénédicte Garnier

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 92

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.