

Module 2 – MODÈLE DE DONNÉES LONGITUDINALES 1

I. OBJECTIFS DU COURS

Ce cours avancé en statistique appliqué aux données hiérarchiques longitudinales a pour but de permettre aux participants d'établir de meilleurs liens entre les données longitudinales et les théories des sciences sociales ou de la santé postulant un effet du temps sur leur objet d'étude (sociologie, psychologie, criminologie, démographie, médecine, etc.). Dans une perspective incluant les variables de contrôle, médiatrices et modératrices ainsi que les concepts relatifs à la causalité, les participants se familiariseront avec les principes fondamentaux nécessaires à la compréhension des techniques d'analyse par trajectoires de développement. Ce cours met l'accent sur les modèles statistiques dont la base est la régression linéaire généralisée avec composantes aléatoires paramétriques.

II. PRÉREQUIS

Ce cours nécessite une connaissance préalable de la statistique descriptive et inférentielle ainsi que des analyses multivariées (régression linéaire simple et multiple). Les connaissances introduites dans le module 1 sont également nécessaires. Plus précisément, la connaissance des modèles spécifiant des liens différentiels en fonction de sous-groupes dans la population (variable modératrice) ou spécifiant des mécanismes liant une variable dépendante à une variable indépendante (variable médiatrice). Il est également important de maîtriser un logiciel statistique comme SPSS, SAS, R, Mplus ou Stata. Les démonstrations de chacune des analyses sont effectuées à partir du logiciel SAS.

Puisque le cours porte sur les concepts et méthodes d'analyse plutôt que sur l'usage d'un logiciel spécifique, il est possible de suivre le cours sans utiliser SAS. Les participants désireux d'utiliser un autre logiciel que SAS seront dirigés vers les ressources utiles pour reproduire les analyses présentées. Ils ne recevront toutefois pas de support en ce qui a trait à l'usage d'un autre logiciel que SAS.

III. DÉMARCHE PÉDAGOGIQUE

L'approche pédagogique de ce cours met l'emphase sur l'écoute de capsules vidéo et la résolution de problèmes en ligne. Les chapitres suivent la séquence suivante: une lecture, une présentation théorique (vidéo), un mini-quiz, un exemple d'application (vidéo), des exercices pratiques et, enfin, un quiz final. Des rencontres synchrones auront lieu à quelques reprises. Le logiciel Adobe Connect sera utilisé lors de ces rencontres.

IV. CONTENU DES CHAPITRES DU COURS

Chapitre 1 : L'exploration des données longitudinales

- Modèle de niveau 1 pour l'analyse du processus de changement individuel
- Modèle de niveau 2 pour l'analyse des différences interindividuelles systématiques dans le changement
- Interprétation des paramètres estimés (effets fixes et effets aléatoires)

Chapitre 2 : Spécification composite du modèle multiniveau du changement

- Spécification des composantes du modèle multiniveau du processus de changement
- Modèle inconditionnel de la moyenne et modèle inconditionnel de croissance
- Développement et adaptation d'une taxonomie de modèles
- Comparaison de modèles

Chapitre 3 : Variabilité dans la durée

- Variabilité de la durée des intervalles de temps en analyse multiniveau
- Variabilité du nombre de temps de mesure
- Avantages et inconvénients de cette flexibilité

Chapitre 4 : Inclusion de variables indépendantes variant dans le temps

- Spécification du modèle incluant des variables variant dans le temps
- Illustration des trajectoires incluant des variables variant dans le temps
- Effet différentiel, par vagues, d'une variable variant dans le temps

Chapitre 5 : Révision de la régression logistique

- Analyse de régression logistique simple
- Fondements et postulats de la régression logistique
- Équations et tests d'hypothèse

Chapitre 6 : Loi de Poisson, loi binomiale négative et distributions modifiées en zéro

- Données comptées
- Loi de Poisson, binomiale négative et les distributions modifiées en zéros
- Généralisation aux modèles multiniveaux

Chapitre 7 : Introduction à l'analyse des données manquantes

- Types de données manquantes et leurs conséquences sur les analyses
- Postulats et mécanismes générant des données manquantes
- Approches contemporaines aux données manquantes
- Imputation multiple, maximum de vraisemblance et pondération