Les équations structurelles

Présentation pour le cours SOL6210, Analyse quantitative avancée

©Claire Durand, 2024

Les équations structurelles

Origine

- Les équations structurelles proviennent de trois “types” d’analyse
  - A) les analyses de cheminement (path analysis)
  - B) les analyses de variables latentes (analyses factorielles)
  - C) les analyses d’équations simultanées
- C’est la combinaison de ces trois *chemins* qui a mené aux équations structurelles:
  - Il s’agit d’une *analyse de cheminement*, comprenant habituellement de *multiples analyses simultanées* et utilisant des *variables latentes*.

Les analyses par équations structurelles sont:

Selon Goodwin (1988)

- Un outil heuristique pour organiser et exprimer la pensée du chercheur concernant la nature attendue des relations entre les variables.
- Un outil permettant de produire des estimations empiriques des relations théoriquement postulées entre des construits.
- Une stratégie pour décider du modèle le plus approprié.

La causalité

Quand on utilise des équations structurelles...

- On présume de l'existence de relations de causalité et donc répondant aux critères suivants:
  - Deux variables covarient.
  - Il y a un ordre (antécédent → conséquence) dans la relation.
  - On a écarté les autres causes possibles de l'effet.
  - L'explication de la chaîne de causalité est possible.

L’importance de la théorie

- “Rien de plus pratique qu’une bonne théorie”. Kurt Lewin
- La complexité doit être modélisée.
- C’est la base des équations structurelles: on cherche à valider empiriquement un modèle théorique. Il faut donc
  - A) avoir élaboré le modèle théorique (model building)
  - B) être en mesure de le tester empiriquement (model testing)

Concrètement,

Les logiciels

- Les premiers logiciels spécifiques:
  - Lisrel de Joreskog et Sorbom (SSICentral);
  - EQS de Bentler
- Les procédures plus récentes, liées à des progiciels
  - Amos (module de SPSS)
  - STATA a une procédure d'équations structurelles avec interface graphique.
  - M plus, R.
- Mon préféré est Lisrel, celui que nous verrons dans le cours mais beaucoup de chercheurs utilisent M+.
- La plupart des logiciels permettent de récupérer les données provenant d'autres logiciels.

Lisrel is greek to me (Hayduk)

- Dès le départ, on a utilisé les équations avec les lettres grecques (référant à la population), particulièrement dans Lisrel.
- Deux séries d'équations
  - A) Le modèle de mesure (équivalent à l'analyse factorielle):
  - $x = \lambda\xi + \delta$ X = lambda ksi+delta

    $y = \lambda\eta + \varepsilon$ Y = lambda eta+epsilon
  - B) Le modèle de relation (équivalent à des régressions simultanées)
  - $\eta = \beta\eta + \gamma\xi + \zeta$ Eta = beta eta+ gamma ksi+ zeta

Quelques notions importantes

Dans Lisrel ou tout autre logiciel

- On fait la différence
  - Entre les variables mesurées (x_i, y_i)
  - Et les variables latentes (ξ(ksi), η (eta))
- Et on fait la différence
  - Entre les variables exogènes, les ξ (ksi), soit les variables qui ne sont pas influencées par aucune autre variable,
  - Et les variables endogènes, les η (eta), qui sont influencées par au moins une autre variable, qu'elle soit exogène ou endogène.

Concrètement,...

Les étapes

- D'abord, théorie, modèle élaboré avant la cueillette de données (pas toujours possible mais c'est l'idéal). On recueille les données en fonction du modèle.
- Après la collecte de données (avant l'analyse d'équations structurelles)
  - Statistiques descriptives univariées.
  - Modèle de mesure: statistiques descriptives multivariées (analyses factorielles exploratoires).
  - Analyses bi-variées et vérification d'hypothèses "simples" (χ^2, tests F, etc.).
  - Régressions avec entrée séquentielle pour avoir une idée de "ce qui tient la route".
  - Élaboration du modèle à tester.

Concrètement, ... suite

Les étapes

- Une fois le modèle élaboré, on est prêt à le tester:
  - A) Faire produire la matrice de covariance ou de corrélation (habituellement nécessaire) et donc déterminer le type de matrice à analyser (peut varier selon le type de variable, le respect des postulats de normalité, etc.).
  - B) Déterminer les paramètres fixes et ceux à estimer, les équations, le programme (langage Simplis dans Lisrel).
  - C) Vérifier l'adéquation du modèle au moyen des divers indices d'ajustement et des indices de signification pour chaque paramètre.
  - D) Réviser le modèle si nécessaire en fonction des indices fournis par Lisrel ET de la théorie à l'origine du modèle.

En conclusion,

- Les équations structurelles constituent un outil très intéressant puisqu'elles permettent de vérifier un ensemble de relations simultanément.
- Elles permettent entre autres de valider des modèles d'analyse factorielle en permettant la présence de variables complexes ou de facteurs de deuxième degré.
- Elles permettent de tester les effets de médiation de façon simultanée.
- On ne peut toutefois pas les utiliser à toutes les sauces. **On doit les garder pour une validation en fin de processus.**