



Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

- Introduction
- Modèle statistique
- Estimation
- Jeu de données

Programmation Mplus

- Pourquoi Mplus?
- Modèle par défaut
- Modèles réalistes
- Modèle avec covariances
- Modèle avec variances flexibles
- Ajout d'un prédicteur de classe

Références

Analyse de classes/profils latents

Concepts et programmation Mplus

Hans Ivers, Ph.D.

École de psychologie, Université Laval

février 2019



Plan de la présentation

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

1 Concepts

- Introduction
- Modèle statistique
- Estimation
- Jeu de données

2 Programmation Mplus

- Pourquoi Mplus?
- Modèle par défaut
- Modèles réalistes
- Modèle avec covariances
- Modèle avec variances flexibles
- Ajout d'un prédicteur de classe

3 Références



Introduction

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Différents noms pour ces techniques dans la littérature

- Latent Class Analysis (LCA) pour des variables nominales ou ordinales
- Latent Profile Analysis (LPA) pour des variables continues
- Model-based Clustering

Toutes ces techniques appartiennent à la famille des modèles de mélange (*mixture modeling*).



Introduction

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

On distingue cette approche (basée sur un modèle) de deux autres approches plus descriptives, restreintes à des données continues, et regroupées sous le vocable d'analyse de regroupement (*cluster analysis*) :

- 1 approche itérative *K-means*, efficace mais non-inférentielle, appropriée pour les sous-populations de "forme similaire" (variances et covariances similaires entre les classes)
- 2 approche hiérarchique (divisive ou agglomérative) basée sur une mesure de distance entre les individus.



Pourquoi cette technique?

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

A partir d'un échantillon d'observations multivariées (chaque répondant est mesuré sur plusieurs variables), l'objectif est d'identifier des sous-populations (classes) homogènes d'individus.

En clair, l'analyste connaît la valeur des variables pour chaque observation, mais ignore à quelle classe chaque observation appartient. On désire donc ici regrouper des observations, contrairement à l'*analyse factorielle* qui regroupe des variables.



Pourquoi cette technique?

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec covariances

Modèle avec variances flexibles

Ajout d'un prédicteur de classe

Références

Exemple univarié d'une LPA à deux classes (Oberski, 2016):

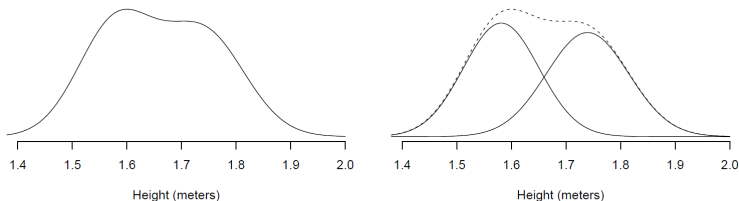


Fig. 1 Peoples' height. Left: observed distribution. Right: men and women separate, with the total shown as a dotted line.

IMPORTANT : le modèle suppose que la forme de la distribution complète des observations correspond à la somme pondérée des distributions de chacune des P sous-populations.



Modèle statistique

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Mathématiquement, pour P classes et k variables :

$$f(Y) = \sum_{i=1}^P \pi_i f_i(Y, \theta_i)$$

avec f la forme de la distribution (on parle de *densité* en statistique), Y est l'ensemble des observations multivariées (une matrice de dimension $n \times k$), π_i le poids de la i ème sous-population (probabilité allant de 0 à 1), et θ_i les paramètres associés à la i ème sous-population (p.ex., la moyenne et la variance pour une distribution normale).



Estimation des paramètres

Analyse de classes/profils latents

Hans Ivers, Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec covariances

Modèle avec variances flexibles

Ajout d'un prédicteur de classe

Références

L'analyste décide :

- du nombre de classes (comme en AF)
- de la forme f des distributions : *normale* pour des variables continues (LPA), *multinominale* pour des variables nominales ou ordinales (LCA)

Le logiciel estime :

- la probabilité de chaque classe π_i
- les paramètres associés à chaque classe θ_i
- le degré d'ajustement du modèle avec les données

Le modèle estimé est donc :

$$f(Y) = \sum_{i=1}^P \hat{\pi}_i f_i(Y, \hat{\theta}_i)$$



Estimation des paramètres

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Statistiquement, c'est un problème de données manquantes (on a les observations mais il manque les classes).

Solution? Utiliser un outil d'estimation pour les données manquantes, soit l'algorithme EM (*Expectation Maximization*).

Approche itérative en trois étapes :

- 1- On choisit au hasard P moyennes/variances (θ_i) multivariées pour les P classes recherchées

Tant que les estimations de θ_i sont trop différentes de l'itération précédente (boucle)

- 2- Placer les observations dans la i e classe la plus probable selon $f_i(Y, \theta_i)$
- 3- Mettre à jour les estimations de π_i et des moyennes/variances (θ_i ; amélioré)



Estimation des paramètres

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Illustration de l'estimation itérative des paramètres θ_i
(moyennes et variances) de chaque classe par l'algorithme EM
(source : Wikipedia).



Jeu de données pour simulation

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Jeu de données composé de $P = 3$ classes comprenant $n = 200$ observations chacune, tirées d'une loi normale bivariée ($k = 2$ variables):

$$\text{Classe 1 : } M_1 = [5, 10], \Sigma_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, r_{12} = 0$$

$$\text{Classe 2 : } M_2 = [10, 5], \Sigma_2 = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}, r_{12} = 0.75$$

$$\text{Classe 3 : } M_3 = [12, 12], \Sigma_3 = \begin{bmatrix} 10 & -2 \\ -2 & 4 \end{bmatrix}, r_{12} = -0.32$$



Jeu de données pour simulation

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

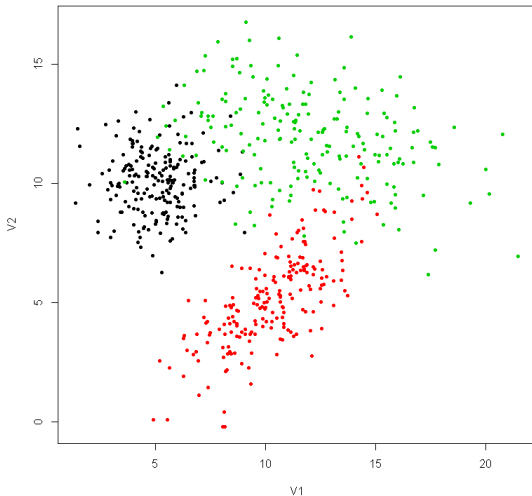
Introduction
Modèle statistique
Estimation
Jeu de données

Programmation Mplus

Pourquoi Mplus?
Modèle par défaut
Modèles réalistes
Modèle avec covariances
Modèle avec variances flexibles
Ajout d'un prédicteur de classe

Références

Illustration en deux dimensions du jeu de données :





Jeu de données pour simulation

Analyse de classes/profils latents

Hans Ivers, Ph.D.

Concepts

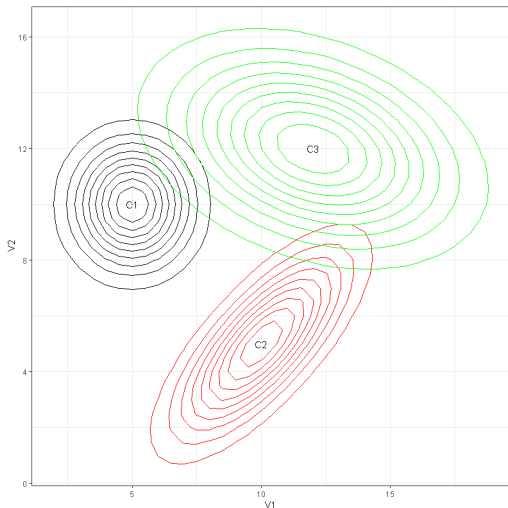
Introduction
Modèle statistique
Estimation
Jeu de données

Programmation Mplus

Pourquoi Mplus?
Modèle par défaut
Modèles réalistes
Modèle avec covariances
Modèle avec variances flexibles
Ajout d'un prédicteur de classe

Références

Densités théoriques pour les trois classes :





Pourquoi Mplus?

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec

covariances

Modèle avec

variances flexibles

Ajout d'un prédicteur

de classe

Références

- Logiciel spécialisé dans les structures latentes (EQS, CFA, LPA/LCA, multiniveaux et autres)
- Développement depuis 1998 par une petite équipe d'experts
- Disponible sous Windows, MacOS et Linux
- Permet de personnaliser chaque modèle
- Offre plusieurs approches d'estimation (maximum de vraisemblance standard ou robuste aux DM, bootstrap, etc.)



Modèle par défaut

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

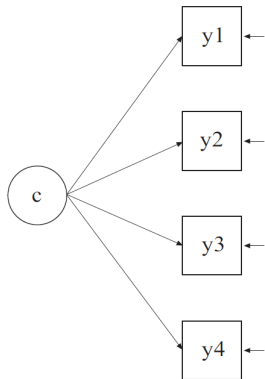
Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

C'est le modèle LPA/LCA le plus simple estimable en Mplus.





Modèle par défaut

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation

Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec

covariances

Modèle avec

variances flexibles

Ajout d'un prédicteur

de classe

Références

Ce modèle a plusieurs postulats :

- Moyennes peuvent varier entre les variables et entre les classes
- Variances peuvent varier entre les variables mais sont fixes entre les classes
- Covariance nulle entre les variables
- Comme la solution est sensible aux valeurs de départ, Mplus génère (par défaut) 20 valeurs aléatoires de θ et vérifie si ajustement est reproduit pour les 4 meilleures. Ce nombre peut être changé avec l'option STARTS.



Modèle par défaut

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Comment programmer une LPA ou une LCA? Par la spécification de la distribution des variables.

Par défaut, une variable est considérée comme continue (i.e., modèle normal) dans Mplus. On doit spécifier le type d'une variable uniquement si elle est non-normale : nominale (NOMINAL), binaire/ordinaire (CATEGORICAL), dénombrement (COUNT) ou de survie (CENSORED).
Donc, pour réaliser une LPA :

```
...  
VARIABLES:  
NAMES ARE id v1-v5;  
...
```



Modèle par défaut

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Pour réaliser une LCA, on devra préciser le type binaire/ordinal des variables (CATEGORICAL):

```
...  
VARIABLES:  
NAMES ARE id v1-v5;  
CATEGORICAL ARE v1-v5;  
...
```

IMPORTANT : la distinction LPA vs LCA est inutile en Mplus car on peut mélanger les deux types de variables dans la même analyse!



Modèle par défaut - syntaxe Mplus

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Syntaxe Mplus pour le modèle par défaut ($P = 3$ classes):

DATA:

```
FILE IS nomdufichier.dat;
```

VARIABLE:

```
NAMES ARE id v1-v2 x1;
```

```
USEVARIABLES ARE v1-v2;
```

```
MISSING ARE a11 (999);
```

```
CLASSES = c(3);
```

ANALYSIS:

```
TYPE = MIXTURE;
```

OUTPUT:

```
TECH1 TECH11 RESIDUAL;
```



Modèle par défaut - nombre de paramètres

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Voici comment calculer le nombre de paramètres à estimer pour un modèle normal (LPA) à P classes et k variables :

- $P \times k$ paramètres pour les moyennes
- k paramètres pour les variances
- 0 paramètre pour les covariances
- $P - 1$ paramètres pour les π_i (probabilité de classe).
Pourquoi $P - 1$? Car $\sum^P \pi_i = 1$.

Exemple : LPA avec 3 classes et 2 variables? Total des paramètres = $3 \times 2 + 2 + 2 = 10$.



Modèle par défaut - taille d'échantillon

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Pas de solution simple..

- Tein et al. (2013): Facteurs les plus importants pour déterminer la puissance du test à déterminer le nombre de classes adéquat: nombre de classes et degré de séparation des classes (et non taille d'échantillon). Un petit échantillon ne pose pas problème si on a peu de classes et qu'elles sont bien distinctes.



Modèle par défaut - taille d'échantillon

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Berg O. Muthén (auteur de Mplus) suggère que :

- “ A common question asked by researchers is, “What sample size do I need for my study?” Over the years, several rules of thumb have been proposed, such as 5 to 10 observations per parameter, 50 observations per variable, no less than 100, and so on. In reality, there is no rule of thumb that applies to all situations. ” (Muthén & Muthén, 2002)
- ”It all depends on how well the classes are separated. I have done successful mixture modeling with only 30 subjects [...]. General rules of thumbs are not worth much for mixtures because results depend so much on the specifics of your situation. That's why we suggest Monte Carlo studies. (Muthén, 2013)



Modèle par défaut - choix du nombre de classes

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec

covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Mplus calcule un certain nombre d'indices pour qualifier l'ajustement du modèle et aider au choix du nombre de classes à conserver pour le modèle final :

- 1 la log-vraisemblance (ne tient pas compte du nombre de paramètres, peu utile)
- 2 les critères d'Akaike (AIC) ou Bayésien (BIC) (une valeur plus faible indique un meilleur ajustement)
- 3 l'indice d'entropie (0 à 1, 1 = parfaite séparation des classes)
- 4 la proportion de chaque classe ($\hat{\pi}_i > 5\%$)



Modèle par défaut - choix du nombre de classes

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

L'examen de ces indices "qualitatifs" peut être complété par des tests statistiques sur le nombre de classes :

- 1 le Lo-Mendell-Rubin (LMR) likelihood ratio test, qui produit une p -value sous H_0 : le modèle avec $P-1$ classes est adéquat (option TECH11)
- 2 le Parametric Bootstrapped likelihood ratio (BLRT) test (option TECH14)

Pour plus de détails, consulter la note de Asparouhov et Muthen (2012), *Using Mplus TECH11 and TECH14 to test the number of latent classes*, disponible en PDF sur le web.



Modèle par défaut - résultats

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction
Modèle statistique
Estimation
Jeu de données

Programmation Mplus

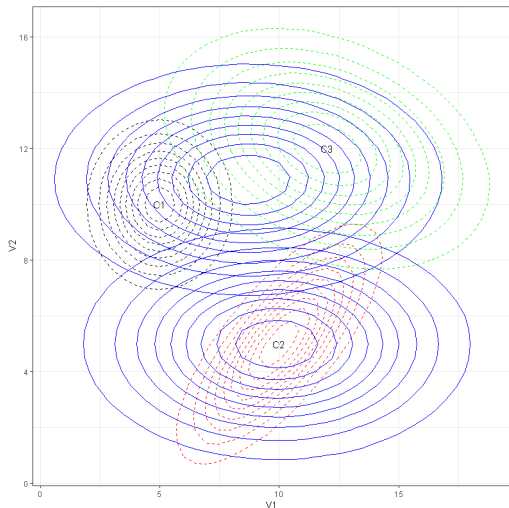
Pourquoi Mplus?
Modèle par défaut

Modèles réalistes
Modèle avec covariances
Modèle avec variances flexibles

Ajout d'un prédicteur de classe

Références

Solution à deux classes :





Modèle par défaut - résultats

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction
Modèle statistique
Estimation
Jeu de données

Programmation Mplus

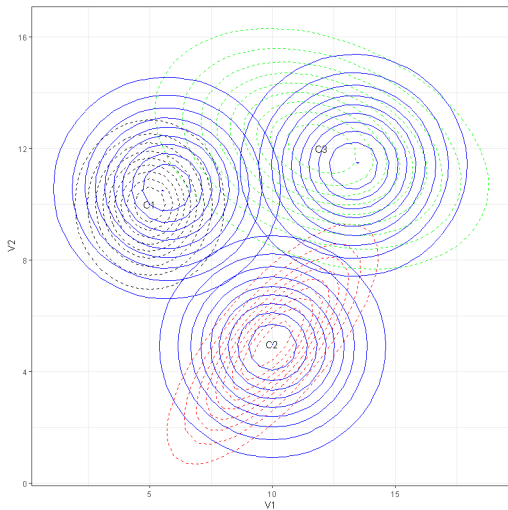
Pourquoi Mplus?
Modèle par défaut

Modèles réalistes
Modèle avec covariances
Modèle avec variances flexibles

Ajout d'un prédicteur de classe

Références

Solution à trois classes :





Modèle par défaut - sauvegarde des classes

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec covariances

Modèle avec variances flexibles

Ajout d'un prédicteur de classe

Références

Option pour sauvegarder les classes assignées et les probabilités a posteriori d'appartenir à chaque classe.

...

OUTPUT:

```
TECH1 TECH11 RESIDUAL;
```

SAVEDATA:

```
FILE IS classes.txt;
```

```
SAVE = CPROBABILITIES;
```



Modèles plus réalistes

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation

Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Le modèle par défaut est simple mais peu réaliste en pratique, en raison de ses nombreuses hypothèses.

L'analyste doit avoir une flexibilité pour préciser :

- qu'il existe des relations entre les variables (les variables ne sont pas toutes parfaitement indépendantes)
- que les variances peuvent être différentes entre les classes (une sous-population peut être plus dispersée qu'une autre sur les variables)
- que certaines variables peuvent prédire la variable latente de classe



Modèle avec covariances

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

- Le modèle par défaut assume une covariance nulle entre chaque paire de variables (hypothèse forte!). Toutefois, cette contrainte permet d'estimer beaucoup moins de paramètres car il y a $k(k - 1)/2$ covariances par classe à estimer.
- *Exemple* ($P = 4$ classes et $k = 6$ variables) : modèle par défaut = $(P + 1)k + P - 1 = 33$ paramètres. Ajout des covariances = $6(5)/2 = 15$ paramètres supplémentaires **PAR CLASSE!**



Modèle avec covariances

Analyse de classes/profils latents

Hans Ivers, Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

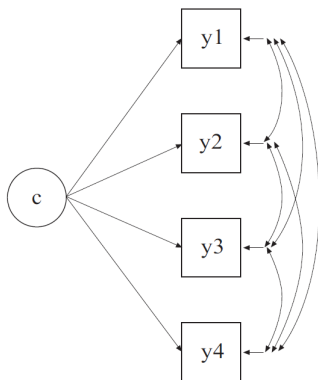
Modèle avec covariances

Modèle avec variances flexibles

Ajout d'un prédicteur de classe

Références

Une approche plus réaliste est de permettre les covariances entre les variables, mais fixer que ces covariances sont les mêmes pour toutes les classes (un postulat connu en MANOVA et en analyse discriminante sous le nom d'*homogénéité des matrices de variance-covariance*).





Modèle avec covariances

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation

Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Addition d'un énoncé MODEL pour ajouter l'estimation des covariances entre les variables (la spécification sous OVERALL force la même valeur pour toutes les classes):

...

ANALYSIS:

```
TYPE = mixture;
```

MODEL:

```
%OVERALL%
```

```
v1 WITH v2;
```

...



Modèle avec covariances - résultats

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

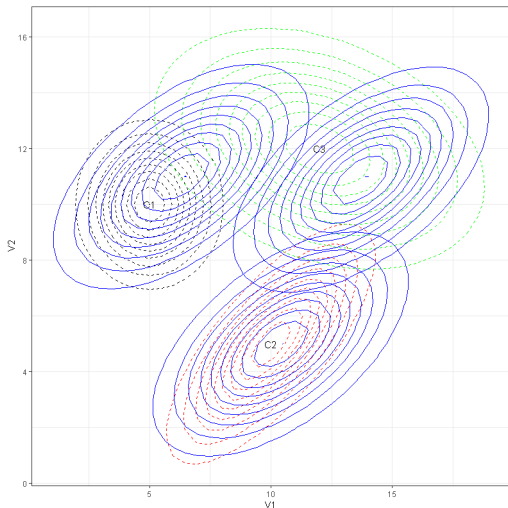
Modèle avec covariances

Modèle avec variances flexibles

Ajout d'un prédicteur de classe

Références

Solution à trois classes :





Modèle avec variances selon classe

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Il est également possible de permettre que les variances de chaque indicateur changent selon la classe (hétérogénéité des variances).

Comme il y a k estimés de variance par classe, cette flexibilité ajoute $(P - 1)k$ paramètres supplémentaires.

Exemple ($P = 4$ classes et $k = 6$ variables) : modèle par défaut = $(P + 1)k + P - 1 = 33$ paramètres. Ajout des variances selon la classe = $(4 - 1)6 = 18$ paramètres supplémentaires (augmentation de 55% de la taille du modèle).



Modèle avec variances selon classe

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation

Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Addition d'un énoncé MODEL pour spécifier d'estimer les moyennes (e.g., [V1]) ET les variances (e.g., V1) par classe :

...

ANALYSIS:

```
TYPE = mixture;
```

MODEL:

```
%c#1%
```

```
[v1 v2];
```

```
v1 v2;
```

```
%c#2%
```

```
[v1 v2];
```

```
v1 v2;
```

```
%c#3%
```

```
[v1 v2];
```

```
v1 v2;
```

...



Modèle avec variances - résultats

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

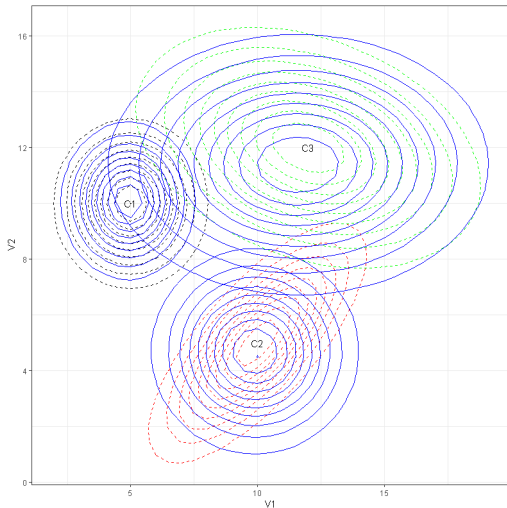
Modèle avec
covariances

**Modèle avec
variances flexibles**

Ajout d'un prédicteur
de classe

Références

Solution à trois classes, avec variances flexibles (selon la classe), covariance nulle:





Modèle avec variances - résultats

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

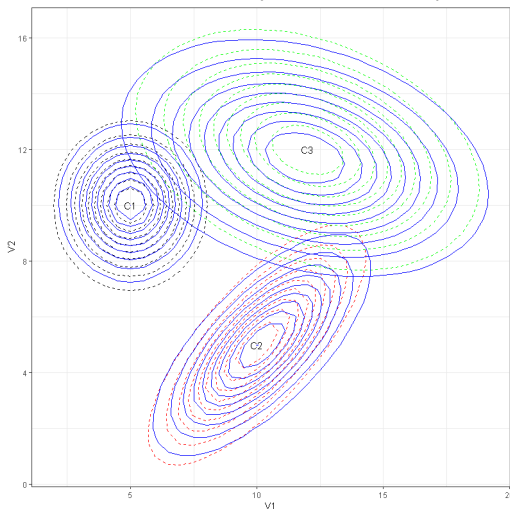
Introduction
Modèle statistique
Estimation
Jeu de données

Programmation Mplus

Pourquoi Mplus?
Modèle par défaut
Modèles réalistes
Modèle avec covariances
Modèle avec variances flexibles
Ajout d'un prédicteur de classe

Références

Solution finale à trois classes, avec variances flexibles et covariances flexibles (selon la classe):





Ajout d'un prédicteur de classe

Analyse de classes/profils latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

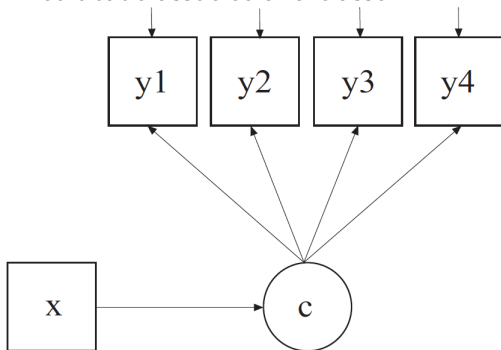
Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

Il est possible d'ajouter des variables qui prédisent la variable latente de classe. Ce modèle est plus avantageux que d'étudier ces relations à l'extérieur de Mplus car on tient compte de l'incertitude associée à la classe.





Ajout d'un prédicteur de classe

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

...

ANALYSIS:

```
TYPE = MIXTURE;
```

MODEL:

```
%OVERALL%
```

```
c ON x1;
```

...

La variable à gauche du ON est prédite (VD, la classe c) alors que celle de droite est le prédicteur (VI, la variable x1).



Quelques références

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec
covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

- Caron, P.-O. (2018). *La Modélisation par Équations Structurelles avec Mplus*. PUQ.
- Collins, L. M., & Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Wiley.
- Muthén, L.K. & Muthén, B.O. (2012). *Mplus User's Guide (7th ed.)* (chapitre 7). Los Angeles, CA.
- Oberski, D.L. (2016). Mixture models : latent profile and latent class analysis. Dans J. Robertson & M. Kaptsein (Eds.) *Modern Statistical Methods for HCI* (pp. 275-287). Springer.



Quelques références (suite)

Analyse de
classes/profils
latents

Hans Ivers,
Ph.D.

Concepts

Introduction

Modèle statistique

Estimation

Jeu de données

Programmation
Mplus

Pourquoi Mplus?

Modèle par défaut

Modèles réalistes

Modèle avec

covariances

Modèle avec
variances flexibles

Ajout d'un prédicteur
de classe

Références

- Muthén BO. *Sample size for LCA*. MPlus 2013; <http://www.statmodel.com/discussion/messages/23/12750.html?1370464379>. Accessed 05-25, 2018.
- Muthén LK, Muthén BO. How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. *Structural Equation Modeling: A Multidisciplinary Journal*. 2002;9(4):599-620.
- Tein J-Y, Coxe S, Cham H. Statistical Power to Detect the Correct Number of Classes in Latent Profile Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*. 2013;20(4):640-657.