WLE en longitudinal 2 temps de mesure

1 Notation

Soit un questionnaire unidimensionnel administré à deux temps de mesure.Les données étant groupées en deux groupes, l'invariance de la mesure est considérée. le partial credit model s'écrit alors :

$$\begin{cases} \mathbb{P}(X_{ijgt} = x|\theta_{igt}, \delta_{j1}, \cdots, \delta_{j\ m_j}) = \dfrac{\mathrm{e}^{x\cdot\theta_{igt}\ -\sum_{p=1}^{x}\delta_{jp}}}{\sum_{l=1}^{m_j}\mathrm{e}^{l\cdot\theta_{igt}-\sum_{p=1}^{l}\delta_{jp}}} \\ \theta_{igt} = \beta_g \cdot g_i + \beta_t \cdot t_i + \beta_{gt} \cdot g_i \cdot t_i + u_i + \theta_{i,residuel} \end{cases}$$

i : individu ; j : item ; t temps de mesure ; θ : niveau du trait latent ; g : groupe ; u_i : effet alétoire (intersept spécifique à l'individu i pour prendre en compte la corrélation individuelle)

2 Méthode

On sait estimer les paramètres de seuils (difficultés) notés $\hat{\delta}_.$, de plus on connaît les réponses des items pour un individu i, les pentes de régression sont estimables notées $\hat{\beta}_.$. On peut aussi estimer a posteriori le vecteur $\hat{U} = \ {}^t(\hat{u_1}; \hat{u_2}; \cdots; \hat{u_n})$.

Sachant ces possibilités, on peut adapter la fonction de vraisemblance proposée par Warm[1, 2] :

$$L(\theta_{igt}|X_{i.gt}) = \prod_{j=1}^{J}\sqrt{I(\theta_{igt})}\cdot\mathbb{P}(X_{ijgt} = x_{ijgt}|\theta_{igt}, \delta_j 1, \cdots, \delta_{j\ m_j})$$

En adaptant avec la formule de régression et les estimations,on cherche à ajuster comme dans une régression linéaire les paramètres θ de sorte à obtenir l'estimation de θ la plus proche possible de la valeur réelle :

$$L(\theta_i|X_{i.gt}, \hat{\beta_g}, \hat{\beta_t}, \hat{\beta_{gt}}, \hat{u_i}) = \prod_{j=1}^{J}\sqrt{I(\theta_i, |, \hat{\beta_g}, \hat{\beta_t}, \hat{\beta_{gt}}, \hat{u_i})}\cdot\frac{\mathrm{e}^{x\cdot(\theta_i+\hat{\beta_g}\cdot g_i+\hat{\beta_t}\cdot t_i+\hat{\beta_{gt}}\cdot g_i\cdot t_i+\hat{u_i})\ -\sum_{p=1}^{x}\delta_{jp}}}{\sum_{l=1}^{m_j}\mathrm{e}^{l\cdot(\theta_i+\hat{\beta_g}\cdot g_i+\hat{\beta_t}\cdot t_i+\hat{\beta_{gt}}\cdot g_i\cdot t_i+\hat{u_i})-\sum_{p=1}^{l}\delta_{jp}}}$$

On peut simplifier la notation, soient

$$D = \begin{pmatrix} \hat{u_1} & g_1 & t_1 & g_1 \cdot t_1 \\ \hat{u_2} & g_2 & t_2 & g_2 \cdot t_2 \\ \vdots & \vdots & \vdots & \vdots \\ \hat{u_n} & g_n & t_n & g_n \cdot t_n \end{pmatrix} \qquad \text{et} \qquad \widehat{\beta} = \begin{pmatrix} 1 \\ \hat{\beta_g} \\ \hat{\beta_t} \\ \hat{\beta_{gt}} \end{pmatrix}$$

Ainsi la vraisemblance pondérée s'écrit :

$$L(\theta_i|X_{i.gt}, \hat{\beta_g}, \hat{\beta_t}, \hat{\beta_{gt}}, \hat{u_i}) = \prod_{j=1}^{J} \sqrt{I(\theta_i|D_i, \widehat{\beta})} \cdot \frac{e^{x\cdot(\theta_i + D_i\widehat{\beta}) - \sum_{p=1}^{x} \delta_{jp}}}{\sum_{l=1}^{m_j} e^{l\cdot(\theta_i + D_i\widehat{\beta}) - \sum_{p=1}^{l} \delta_{jp}}}$$

L'estimateur du trait latent de Warm adapté dans un cadre longitudinal serait :

$$\widehat{\theta_{igt}} = \underset{\theta_i \in \mathbb{R}}{\text{argmax}}\ L(\theta_i|X_{i.gt}, \hat{\beta_g}, \hat{\beta_t}, \hat{\beta_{gt}}, \hat{u_i})$$

Or dans ce cas on estime le trait latent résiduel $\theta_{residuel}$ pour obtenir le vraie niveaux de trait latent pour des individus situés dans le groupe g, on doit a posteriori rajouter le terme $D_i \cdot \widehat{\beta}$

$$\widehat{\theta_{igt}} = \left[\underset{\theta_i \in \mathbb{R}}{\text{argmax}}\ L(\theta_i|X_{i.gt}, \hat{\beta_g}, \hat{\beta_t}, \hat{\beta_{gt}}, \hat{u_i}) \right] + D_i \cdot \widehat{\beta}$$

3 Applications

3.1 Cadre transversal

En se restreignant à la variable de groupe (variable exogène), il ne reste plus que $\hat{\beta_g}$ et g_i Ainsi la vraisemblance s'écrit :

$$L(\theta_i|X_{i.g}, \hat{\beta_g},) = \prod_{j=1}^{J} \sqrt{I(\theta_i|g_i, \hat{\beta_t})} \cdot \frac{e^{x\cdot(\theta_i + g_i\hat{\beta_t}) - \sum_{p=1}^{x} \delta_{jp}}}{\sum_{l=1}^{m_j} e^{l\cdot(\theta_i + g_i\hat{\beta_t}) - \sum_{p=1}^{l} \delta_{jp}}}$$

L'estimateur s'écrit alors :

$$\widehat{\theta_{ig}} = \left[\underset{\theta_i \in \mathbb{R}}{\text{argmax}}\ L(\theta_i|X_{i.g}, \hat{\beta_g}) \right] + g_i \cdot \hat{\beta_g}$$

3.1.1 R

Dans R on peut calculer l'estimateur de Warm en utilisant deux packages

1. TAM : en utilisant la fonction IRT.residuals(model [1])$theta
2. PP : en utilisant la fonction PP_gpcm, prenant en entrée les paramètres de seuils estimés et les réponses des items

1. model désigne qu'un modèle PCM a été précédemment estimer avec la fonction tam.mml

3.1.2 STATA®

Dans STATA®, il n'existe pas de fonctions ou de modules natifs pour estimer le trait latent de cette façon. Je propose un module `.ado` qui permet de calculer les ”WLE”. Il prend en entrée

— thres : une matrice des paramètres de seuils estimés

— resp : une chaîne de caractère correspondant aux noms des variables de réponse aux items

— grp : un chaîne de caractère correspondant au nom de la variable de groupe

— beta : la valeur de l'effet groupe estimée

Ce module est disponible sur gitHub Psychometric-Tools-STATA, il est toujours en cours de développement dans le dossier `ado` ; nom du module `wle`.

Dans les deux languages de programmation statistiques les résultats sont identiques à 10^{-6} près ; de plus les temps d'exécution sont relativement proches.

3.2 Cadre longitudinal

3.2.1 R

Pour l'instant il n'existe pas de package à ma connaissance permettant d'estimer un PCM longitudinal. On ne peut donc pas implémenter quoi que se soit en lien avec les estimations des paramètres d'un tel modèle $\rightarrow$ impossible d'appliquer la méthode.

3.2.2 STATA®

Avec le module `gsem` on a la possibilité d'estimer tout les paramètres d'un PCM longitudinal nécessaire à l'application de la méthode. Je propose de la même façon un module `.ado` qui prends en entrée les mêmes paramètre que le module transversal. En plus de ça, il y a 3 paramètre supplémentaire :

— U le nom de la variable des estimations de l'effet aléatoire pour chaque individu.

— $time$ le nom de la variable correspondant au temps

— $beta$ n'est plus un scalaire mais un vecteur colonne qui contient 1, puis les paramètres de régression latente estimés

4 Autre possibilité extension : à N-temps de mesure

Comme on vient de le voir, on a la possibilité d'estimer le trait latent pour un individu à deux temps de mesures. L'objectif majeur est de regarder l'évolution du trait latent au cours du temps. La méthode des WLE a été étendue vers un méthode plus générale en utilisant des modèles espace-états [3]. De plus, une autre méthode a été présentée pour suivre l'évolution du trait latent dans le temps (temps discret), en utilisant de la même manière des modèles espace-états [4]. Cette méthode pourrait être utile si les questionnaire utilisés sont régulièrement passés.

L'avantage de ces modèles et de leurs applications dans l'IRT c'est de pouvoir faire de la prédiction du trait latent sur une longue ou courte période de temps.

Une application de ces modèles pourraient être de suivre par exemple l'évolution moyenne de l'anxiété pour des personnes atteintes de dépression longue. Pour étudier par exemple des phénomènes de saisonnalité ou autres. Mais aussi pour comparer sur le long terme l'impact d'une intervention ou d'un traitement.
Cette section est toujours en cours de réflexion et de développement.

Références

1. Warm TA. Weighted Likelihood Estimation of Ability in Item Response Theory. Psychometrika. 1989 Sep; 54 :427-50. doi : 10.1007/BF02294627. Available from : https://www.cambridge.org/core/product/identifier/S0033312300030416/type/journal_article [Accessed on : 2025 May 5]
2. Penfield RD et Bergeron JM. Applying a Weighted Maximum Likelihood Latent Trait Estimator to the Generalized Partial Credit Model. Applied Psychological Measurement. 2005 May; 29 :218-33. doi : 10.1177/0146621604270412. Available from : https://journals.sagepub.com/doi/10.1177/0146621604270412 [Accessed on : 2025 May 5]
3. Kamphuis F. Estimation and Prediction of Individual Ability in Longitudinal Studies. Instituut voor Toetsontwikkeling, 1992
4. Wanjohi RG, Rijn PW van et Davier AA von. A State Space Approach to Modeling IRT and Population Parameters from a Long Series of Test Administrations. *New Developments in Quantitative Psychology*. Sous la dir. de Millsap RE, Ark LA van der, Bolt DM et Woods CM. New York, NY : Springer, 2013 :115-32. doi : 10.1007/978-1-4614-9348-8_8