The latent factor structure of child development

Anonymous Cogsci Submission

Abstract

Hello

Keywords: one; two;

Introduction

A child's development can be thought of as the set of developmental milestones that they have reached at a particular point in time. This conceptualization results in data with the same structure as the item response data common to educational measurement. In education, item response data is most typically students responding to test items (i.e., questions) and, in the dichotomous case, getting each question either correct or incorrect. In the context of child development, the child is the "student," and each developmental milestone is the "item."

We use Kinedu, a Mexico-based child development app, as a source for this type of data. When parents first start using the Kinedu app, they are asked a series of questions about which developmental milestones their child has reached. We consider the 1946 children between 2 and 55 months of age whose parents responded to all 414 of the developmental milestones. Each developmental miletone on Kinedu is mapped to a milestone group: physical, cognitive, linguistic, or social & emotional. Table 1 shows the number of developmental milestones in each group along with an example milestone in Spanish (as its shown to the parent) and its translation in English.

Figure 1 shows the age (in months) and number of developmental milestones for each child. As can be seen in Figure 1, at 12 months of age, most children have reached about 200 developmental milestones. At 24 months of age, most children have reached about 300 developmental milestones. Finally, at 48 months of age, most children have reached about 375 of the 414 developmental milestones.

Empirical assessment of the dimensionality of child development

We frame the assessment of the dimensionality of child development as a model comparison question.

Models

Item response theory offers a suite of models with which to model item response data. We adopt the notation used in Chalmers & others (2012). Let $i = 1, \ldots, I$ represent the distinct children and $j = 1, \ldots, J$ the developmental milestones. The Kinedu item response data is stored in a matrix, y, where element y_{ij} denotes if the ith child has or has not achieved the jth developmental milestone as reported by their parent/guardian. Each model represents the ith child's development using m latent factors $\boldsymbol{\theta}_i = (\theta_1, \ldots, \theta_m)$. The jth milestone's discriminations (i.e. slopes) $\boldsymbol{a_j} = (a_1, \ldots, a_m)$ capture the latent factor loadings onto that milestone.

We fit four two-parametric logistic (2PL) models where a child's development is represented by $m = 1, m = 2, m = 3$, and $m = 4$ latent factors. According to the 2PL model, the probability of a child having achieved a developmental milestone is

$$P(y_{ij} = 1 | \boldsymbol{\theta_i}, \boldsymbol{a_j}, b_j) = \sigma(\boldsymbol{a}_j^\top \boldsymbol{\theta_i} + b_j)$$

where b_j is the milestone easiness (i.e. intercept) and $\sigma(x) = \frac{e^x}{e^x+1}$ is the standard logistic function.

The 2PL models learn the latent factor structure entirely from the data, making them exploratory. The bifactor model offers an alternative specification where each milestone loads onto a general factor θ_0 and a specific factor θ_s. The assignment of each developmental milestone to its specific factor is an opportunity to specify the latent factor structure, making the model confirmatory as opposed to exploratory. We map each milestone to its specific factor according to the four developmental milestone groups shown in Table X. For the bifactor model, the probability of a child having achieved a developmental milestone is

$$P(y_{ij} = 1 | \theta_0, \theta_s, a_0, a_s) = \sigma(a_0\theta_0 + a_s\theta_s + b_j).$$

Model comparison

Model comparison in IRT typically uses information criterion such as AIC and BIC (Maydeu-Olivares, 2013). However, these methods are not guaranteed to work at modest sample sizes (McDonald & Mok, 1995). Instead, we prefer a marginalized version of cross-validation. In essence, we partition the data into folds based on the children (i.e. the rows of the item response matrix). Then for each fold, we estimate the item parameters using all but that fold, and calculate the likelihood of that fold by integrating over $g(\theta)$.

Mathematically and following notation similar to Vehtari, Gelman, & Gabry (2017), we partition the data into K subsets

$y^{(k)}$ for $k = 1, \ldots, K$. Each model is fit separately to each training set $y^{(-k)}$ yielding item parameter estimates which we compactly denote $\mathbf{\Psi}^{(-k)}$. The predictive (i.e. out-of-sample) likelihood of $y^{(k)}$ is

$$p(y^{(k)}|y^{(-k)}) = \prod_{i \in i^{(k)}}^{I} \int_{\theta} \prod_{j=1}^{J} \hat{\Pr}(y_{ij}^{(k)}|\Psi_j^{(-k)}, \theta) g(\theta) d\theta.$$

The ultimate quantity of interest for each model is the log predictive likelihood for the entire item response matrix, which is defined as

$$\text{lpl } y = \sum_{k=1}^{K} \log p(y^{(k)}|y^{(-k)}).$$

Results
HELLO

Acknowledgements

We'd like to thank Kinedu for providing the data that made this research possible.

References

Chalmers, R. P., & others. (2012). Mirt: A multidimensional item response theory package for the r environment. *Journal of Statistical Software*, *48*(6), 1–29.

Maydeu-Olivares, A. (2013). Goodness-of-fit assessment of item response theory models. *Measurement: Interdisciplinary Research and Perspectives*, *11*(3), 71–101.

McDonald, R. P., & Mok, M. M.-C. (1995). Goodness of fit in item response models. *Multivariate Behavioral Research*, *30*(1), 23–40.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, *27*(5), 1413–1432.