An Analysis of Clinical Data: Illustrating Equivalence of Unidimensional Item Response Theory and Cognitive Diagnosis Models

Kevin Carl P. Santos, Ph.D.
School of Statistics
University of the Philippines-Diliman

Jimmy de la Torre, Ph.D.
Faculty of Education
The University of Hong Kong
Introduction

• At present, many existing educational assessments are developed and analyzed using unidimensional item response theory (IRT) models, which assume a single continuous latent variable.

• To extract more diagnostic information, the same assessments have been retrofitted with cognitive diagnosis models (CDMs), which assume a multidimensional discrete latent variable.

• However, it is not clear to what extent disparate psychometric frameworks can be used on the same data.

• To address this issue, we propose a unifying framework for relating the two classes of model, as well as boundaries as to when this can be done.
Unidimensional IRT Models

Proportional Reasoning

Ability Level $\theta$

Low Ability

High Ability
Cognitive Diagnosis Models

Mastery Probability

Non-mastered

Mastered

Attributes in Proportional Reasoning

- Prerequisite skills
- Comparing fractions
- Ordering fractions
- Constructing ratios
- Constructing proportions
- Multiplicative relationship
- Proportional relationship
- Applying algorithm

Mastery Probability

0% 20% 40% 60% 80% 100%
Equivalence of IRT and CDM

• In CDM, the marginal probability of $x_j$ can then be written as

$$p(x_j) = \sum_{l=1}^{L} p(x_j|\alpha_l)p(\alpha_l),$$

where $p(x_j|\alpha_l)$ is the *item response model* and $p(\alpha_l)$ the *joint attribute distribution*.

• One way of specifying $p(\alpha_l)$ is to use a unidimensional higher-order latent trait, such that

$$p(\alpha_l|\theta) = \prod_{k=1}^{K} p_k(\alpha_{lk}|\theta),$$

where $p_k(\alpha_{lk}|\theta)$ is the *attribute mastery function* (AMF).
Equivalence of IRT and CDM

• Hence,

$$p(x_j) = \sum_{l=1}^{L} \int_{\theta} p(x_j|\alpha_l)p(\alpha_l|\theta)p(\theta)d\theta$$

• For greater generality, the CDM can be represented by the generalized deterministic inputs, noisy, “and” gate (G-DINA; de la Torre, 2011) model:

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*} \delta_{jkk'} \alpha_{lk} \alpha_{lk'} + \cdots + \delta_{j12\ldots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk},$$

where \(\delta_{j0}\) is the baseline probability, \(\delta_{jk}\)s the main effects, \(\delta_{jkk}\)s the two-way interaction effects, and \(\delta_{j12\ldots K_j^*}\) the highest order interaction effect.
Equivalence of IRT and CDM

- To compare unidimensional IRT models and CDMs, we can express the CDM success probability on item $j$ as a function of $\theta$, as in,

$$p(x_j|\theta) = \sum_{l=1}^{2^K} p(x_j, \alpha_l|\theta) = \sum_{l=1}^{2^{K_j^*}} p(x_j|\alpha_{ij}^*) p(\alpha_{ij}^*|\theta)$$

- We need to further re-write $p(x_j|\theta)$ to better understand its properties.
- For notational convenience, we write $p(\alpha_k=1 | \theta) = p_k(1 | \theta)$ as $p_k$.
- When only one attribute is required, $p(x_j|\theta)$ simplifies to

$$p(x|\theta) = \sum_{\alpha_1=0}^{1} p(x|\alpha_1) p_1(\alpha_1|\theta)$$

$$= \delta_0 + \delta_1 p_1.$$
Equivalence of IRT and CDM

- When $K_j^*$ attributes are required, $p(x_j|\theta)$ can be expressed as

\[
p(x_j|\theta) = \sum_{\alpha_1=0}^{1} \cdots \sum_{\alpha_{K_j^*}=0}^{1} p(x_j|\alpha_1, \ldots, \alpha_{K_j^*}) p(\alpha_1, \ldots, \alpha_{K_j^*}|\theta)
\]

\[
= \delta_0 + \sum_{k=1}^{K_j^*} \delta_k p_k + \sum_{k=1}^{K_j^*-1} \sum_{k'=k+1}^{K_j^*} \delta_{kk'} p_k p_{k'} + \cdots + \delta_{1\ldots K_j^*} \prod_{k=1}^{K_j^*} p_k
\]

- We refer to this as the reformulated HO-GDINA (RHO-GDINA) model
Equivalence of IRT and CDM

Sufficient Conditions for Monotonically Nondecreasing \( p(x_j|\theta) \)

- For \( p(x_j|\theta) \) to be monotonically nondecreasing, the following sufficient conditions need to be met:

1. The AMF of \( p_k, k = 1, \ldots, K \) should be of the form

\[
p_k = \frac{\exp[\beta_k(\theta - \lambda_k)]}{1 + \exp[\beta_k(\theta - \lambda_k)]},
\]

where \( \beta_k \) and \( \lambda_k \) are the discrimination and difficulty parameters with respect to attribute \( k \).

2. \( p(x_j|\alpha_l^*) \leq p(x_j|\alpha_l'\cdot^*) \) whenever \( \alpha_l^* \leq \alpha_l'\cdot^* \) (monotonicity property)
Millon Clinical Multiaxial Inventory-III

- For MCMI-III has been indicated by clinicians as being one of the most frequently used self-report instruments for clinical assessment
- To illustrate the IRT and CDM equivalence, we analyzed the responses of 1,210 subjects to 130 statements of the Dutch version of the MCMI-III
- The statements measure 16 clinical disorders, namely,

<table>
<thead>
<tr>
<th>$\alpha_i$</th>
<th>Disorder</th>
<th>$\alpha_j$</th>
<th>Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$</td>
<td>Depressive</td>
<td>$\alpha_9$</td>
<td>Somatoform</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>Sadistic</td>
<td>$\alpha_{10}$</td>
<td>Bipolar</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>Negativistic</td>
<td>$\alpha_{11}$</td>
<td>Dysthymia</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>Masochistic</td>
<td>$\alpha_{12}$</td>
<td>Drug Dependence</td>
</tr>
<tr>
<td>$\alpha_5$</td>
<td>Schizotypal</td>
<td>$\alpha_{13}$</td>
<td>Post Traumatic Stress</td>
</tr>
<tr>
<td>$\alpha_6$</td>
<td>Borderline</td>
<td>$\alpha_{14}$</td>
<td>Thought Disorder</td>
</tr>
<tr>
<td>$\alpha_7$</td>
<td>Paranoid</td>
<td>$\alpha_{15}$</td>
<td>Major Depression</td>
</tr>
<tr>
<td>$\alpha_8$</td>
<td>Anxiety</td>
<td>$\alpha_{16}$</td>
<td>Delusional Disorder</td>
</tr>
</tbody>
</table>
Millon Clinical Multiaxial Inventory-III: Method

- The model fit of the saturated and higher-order (1-parameter logistic or 1PL and 2PL) G-DINA models, and the four unidimensional IRT models (i.e., 4PL, 3PL, 2PL, and 1PL) were compared

- The AIC and BIC were employed for relative fit evaluation

- The correlations of the different \( \hat{\theta} \)s were calculated

- To compare the IRT and CDM estimates, the number of disorders were plotted against the latent trait estimates
Millon Clinical Multiaxial Inventory-III: Results

### Table: Model Comparison

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>4PL</td>
<td>161130</td>
<td>163781</td>
</tr>
<tr>
<td>3PL</td>
<td>159581</td>
<td>161569</td>
</tr>
<tr>
<td>2PL</td>
<td>159433</td>
<td>160759</td>
</tr>
<tr>
<td>1PL</td>
<td>165675</td>
<td>166343</td>
</tr>
<tr>
<td>Saturated</td>
<td>282580</td>
<td>621740</td>
</tr>
<tr>
<td>2PL-GDINA</td>
<td>155332</td>
<td>160533</td>
</tr>
<tr>
<td>1PL-GDINA</td>
<td>156154</td>
<td>161278</td>
</tr>
</tbody>
</table>

- Among the IRT models, the 2PL obtained the lowest AIC and BIC.
- Based on AIC and BIC, the 2PL-GDINA model fitted the data better than 1PL-GDINA or saturated GDINA model.
- It can be noted as well that when compared with the four IRT models the 2PL-GDINA model had the lowest AIC and BIC.
Millon Clinical Multiaxial Inventory-III: Results

<table>
<thead>
<tr>
<th>Correlation</th>
<th>2PL-GDINA</th>
<th>1PL-GDINA</th>
</tr>
</thead>
<tbody>
<tr>
<td>4PL</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>3PL</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>2PL</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>1PL</td>
<td>0.95</td>
<td>0.97</td>
</tr>
</tbody>
</table>

- The latent trait estimates obtained from the HO-GDINA and the unidimensional IRT models are highly correlated.
- This is consistent with the results of the simulation study and real data analysis on proportional reasoning assessment previously conducted by the authors.
Millon Clinical Multiaxial Inventory-III: Results

<table>
<thead>
<tr>
<th>Disorder</th>
<th>$\beta_k$</th>
<th>$\lambda_k$</th>
<th>Disorder</th>
<th>$\beta_k$</th>
<th>$\lambda_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$</td>
<td>4.29</td>
<td>-0.36</td>
<td>$\alpha_9$</td>
<td>2.41</td>
<td>-0.27</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1.05</td>
<td>-0.74</td>
<td>$\alpha_{10}$</td>
<td>1.31</td>
<td>-0.32</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>3.19</td>
<td>0.46</td>
<td>$\alpha_{11}$</td>
<td>3.40</td>
<td>0.40</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>4.67</td>
<td>-0.90</td>
<td>$\alpha_{12}$</td>
<td>0.10</td>
<td>-1.26</td>
</tr>
<tr>
<td>$\alpha_5$</td>
<td>3.49</td>
<td>-1.23</td>
<td>$\alpha_{13}$</td>
<td>1.72</td>
<td>-0.46</td>
</tr>
<tr>
<td>$\alpha_6$</td>
<td>2.79</td>
<td>-0.98</td>
<td>$\alpha_{14}$</td>
<td>4.12</td>
<td>0.35</td>
</tr>
<tr>
<td>$\alpha_7$</td>
<td>1.48</td>
<td>0.00</td>
<td>$\alpha_{15}$</td>
<td>2.75</td>
<td>-0.29</td>
</tr>
<tr>
<td>$\alpha_8$</td>
<td>4.24</td>
<td>0.31</td>
<td>$\alpha_{16}$</td>
<td>1.58</td>
<td>-1.29</td>
</tr>
</tbody>
</table>

- Except for drug dependence, all slope coefficients had very values, suggesting a strong relationship of each disorder to the general latent trait (i.e., an indication of unidimensionality)
Millon Clinical Multiaxial Inventory-III: Results

Higher Order $\hat{\theta}$ versus Number of Disorders
Millon Clinical Multiaxial Inventory-III: Results

IRT $\hat{\theta}$

versus

Number of Disorders
Millon Clinical Multiaxial Inventory-III: Discussion

• The subjects with higher latent traits possess more disorders

• When the number of disorders was fixed, the corresponding values of the latent trait varied and overlapped with different number of disorders

• Hence, using the latent trait estimate solely would be insufficient in targeting the specific disorders that need to be addressed

• Nevertheless, this work provides a framework for relating the two classes of psychometric models
Millon Clinical Multiaxial Inventory-III: Discussion

• Under certain conditions (e.g., AMF slope is large), the HO-GDINA model can be approximated by IRT models

• As shown in the real data analysis, almost all slope coefficients were very large, producing high correlation of the HO-GDINA and IRT model latent trait estimates

• Thus, in addition to finer-grained attributes, estimating the overall ability (or general latent trait) is also reasonable

• When the HO assumption is reasonable, IRT models can be fitted to CDM data to obtain an approximation of the HO ability
Fin.

Thank you very much!

kpsantos1@up.edu.ph