

When Simpler is Better: Impact of Appropriate Model Selection on Attribute Classification

Wenchao Ma Charles Iaconangelo Jimmy de la Torre
Rutgers, The State University of New Jersey

I. Motivation

1. A variety of cognitive diagnosis models (CDMs) have been developed recently. Determining which model is the most appropriate for an item is not always apparent.

	General or saturated CDMs	Specific or reduced CDMs
Examples	<ul style="list-style-type: none">• G-DINA model (de la Torre, 2011)• General Diagnostic Model (von Davier, 2008)• Log-linear CDM (Henson, Templin, & Willse, 2009)	<ul style="list-style-type: none">• DINA model (Haertel, 1989)• DINO model (Templin & Henson, 2006)• <i>Additive</i> CDM (<i>A</i>-CDM; de la Torre, 2011)• Linear Logistic Model (LLM; Maris, 1999)• Reduced Reparametrized Unified Model (R-RUM; Hartz, 2002)
Advantages	<ul style="list-style-type: none">• Always provide better model-data fit• Are less likely to cause misfit	<ul style="list-style-type: none">• Are more stable• Have more straightforward interpretations• Provide more accurate classification

2. Bases for model selection

a) Expert judgment

b) The Wald test (de la Torre & Lee, 2013): Evaluating, item by item, whether a saturated model (i.e., G-DINA) can be replaced by a reduced model without a significant loss to model-data fit

3. Considering that the true model is never known, it is not clear whether selected CDMs based on the Wald test can provide comparable classification rates as the true model.

II. Goal

To evaluate whether the attribute classification of CDMs selected based on the Wald test is better than that of general CDMs

III. Study design

Factors

- | | |
|--|---|
| <ul style="list-style-type: none">• Sample size: 500, 1000, 2000• Test length: 30• Number of attributes: 5• Item Quality: High, Moderate, Low | <ul style="list-style-type: none">• Attribute distribution: Uniform, Higher-order• Generating models: DINA, DINO, <i>A</i>-CDM, LLM, R-RUM• Estimating models: True models, selected CDMs based on the Wald test, saturated model (i.e. G-DINA) |
|--|---|

IV. Preliminary results

The CDMs selected by the Wald test have higher correct classification rates than general CDMs, especially when the sample size is small, or the item quality is low.

Authors

Wenchao Ma is a second year Ph.D. student in Educational Measurement and Statistics at Rutgers University. His research interests include cognitive diagnosis modeling, item response theory and their applications in psychological and educational measurement. wenchao.ma@rutgers.edu

Charles Iaconangelo is a second year Ph.D. student in Educational Measurement and Statistics at Rutgers University. His research interests include cognitive diagnosis models and their application to the classroom. charles.iaconangelo@rutgers.edu.