Evaluating Bayesian Transition Diagnostic Classification Models for Reporting Within-Year Progress

Jeffrey C. Hoover & W. Jake Thompson
Importance of Reporting Within-Year Progress

• Supplements performance results by providing additional information to students and parents
• Provides feedback to educators and administrators
• Supports the theory of action for assessments when it involves making progress
Diagnostic Modeling

• Diagnostic classification models (DCMs) assume discrete latent constructs (i.e., attributes)
  • For DCMs, the attributes are frequently binary and labeled as masters and nonmasters
• DCMs estimate the probability that each examinee is a member of each latent class
  • Outputs attribute mastery profiles
Log-Linear Cognitive Diagnosis Models (LCDMs)

• One of the more prevalent DCMs
• Uses an approach similar to ANOVA
  • Measurement model sums the log-odds for the mastered attributes
Transition Diagnostic Classification Models (TDCMs)

• The longitudinal extension of the LCDM
  • The TDCM uses the LCDM measurement model with latent transition analysis
• Models changes in attribute mastery statuses over time
• Item invariance is assumed across assessment points
  • E.g., items are just as difficult at Time 2 as at Time 1
Objectives

• Compare TDCM-based estimates of within-year progress to LCDM-based estimates of within-year progress in a simulation study
  • TDCM
  • Full-year LCDM (separately scoring data from each window)
  • Window-specific LCDMs
## Simulation Factors

<table>
<thead>
<tr>
<th>Factor/Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition from mastery to nonmastery</td>
<td></td>
</tr>
<tr>
<td>Unconstrained</td>
<td>$U[0.00, 1.00]$</td>
</tr>
<tr>
<td>Moderate constraint</td>
<td>$U[0.00, 0.50]$</td>
</tr>
<tr>
<td>Large constraint</td>
<td>$U[0.00, 0.15]$</td>
</tr>
</tbody>
</table>

Accessible Teaching, Learning, & Assessment Systems
Data Structures

• We simulated the data based on data collected from an operational alternate assessment from 2016—2017 to 2021—2022
  • Assessment is intended to be scaled with a DCM
  • Skills are individually modeled using single-attribute LCDMs
    • Produces TDCMs with 4 possible transitions
Simulated Parameters

• We based the item parameters and base rate of mastery in each repetition on randomly selected models from the alternate assessment’s operational calibration
  ○ Produces operationally realistic parameter values

• The items in the alternate assessment are assumed to be fungible
Example Transition Matrix

<table>
<thead>
<tr>
<th>Fall</th>
<th>Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonmaster</td>
</tr>
<tr>
<td>Nonmaster</td>
<td>.30</td>
</tr>
<tr>
<td>Master</td>
<td>.20</td>
</tr>
</tbody>
</table>
Data Simulation

• Simulate number of examinees and items based on data structure
• Establish true parameter values
• Assign true transitions to students
• Simulate item responses based on true transition and parameter values
Model Evaluation

• Classification accuracy
  • Defined as the percent correct

• Measured at two levels
  • Overall classification accuracy (student-level transitions)
  • Marginal classification accuracy (student-level mastery in the fall and spring)
Model Estimation Results

• 900 estimated TDCMs
• 2,566 estimated LCDMs
  • 872 (97%) full-year LCDMs
  • 1,694 (94%) window-specific LCDMs
• All 134 LCDMs that did not complete took longer than 12 hours to estimate
## Classification Accuracy

<table>
<thead>
<tr>
<th>Type of classification accuracy</th>
<th>Transition constraint</th>
<th>TDCM</th>
<th>Full-year LCDM</th>
<th>Window-specific LCDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>Unconstrained</td>
<td>.80</td>
<td>.60</td>
<td>.66</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>.78</td>
<td>.58</td>
<td>.61</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>.78</td>
<td>.63</td>
<td>.65</td>
</tr>
<tr>
<td>Marginal – Fall</td>
<td>Unconstrained</td>
<td>.88</td>
<td>.74</td>
<td>.77</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>.87</td>
<td>.70</td>
<td>.72</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>.86</td>
<td>.70</td>
<td>.72</td>
</tr>
<tr>
<td>Marginal – Spring</td>
<td>Unconstrained</td>
<td>.89</td>
<td>.78</td>
<td>.82</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>.88</td>
<td>.77</td>
<td>.79</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>.88</td>
<td>.81</td>
<td>.83</td>
</tr>
</tbody>
</table>
Summary of Results

• The TDCM showed higher classification accuracy than the LCDM-based approaches
• Classification accuracies were consistent across the transition constraint
Discussion

• LCDM-based approaches appeared to miss significant aspects of within-year progress
• Full-year LCDM aggregates data across windows
  • Changes in attribute mastery may be obscured
• Window-specific LCDM did not assume item invariance
  • Progress as evidenced by improved performance may be interpreted as easier items
Thank you!

Contact me at: jhoover4@ku.edu