A Step Further with Computerised Adaptive Testing

The 'Snapshot' Assessment

Charles Darr and Hilary Ferral
NZCER
9 September 2011
Overview

• Computer assisted testing and the Literacy and Numeracy for Adults Assessment Tool
• Motivation for the snapshot assessment
  – What do we want to be able to do?
• Ingredients of the snapshot assessment
  – Psychometric model
  – Item bank(s)
  – The assessment
    • Starting point for assessment
    • Item selection algorithm
    • Termination criteria
CAT Methodology – A Learner Story

9 September 2011
Symposium on Assessment and Learner Outcomes
Another CAT Learner Story
Motivation for the Snapshot

- Short CAT
- Classification
  - Above line
  - Below line

Numeracy Scale
Psychometric Model

• We are not starting with a blank slate
• Rasch model
  – Probability of success on an item
• Large bank of items which probe skills in particular areas of learning
• Existing literacy and numeracy scales
  – Scales developed with original tool
• Existing cut-points describing “step levels”
Building a Snapshot Assessment

• Need to establish 3 things
  – How to start
    • Item difficulty
    • Learner’s starting point
  – How to proceed in the middle
    • Item selection algorithm
  – How to finish
    • Categorise learner?
Starting the Snapshot Assessment

• How difficult should the first item(s) be?
• Options
  – Start in the middle of the bank?
  – Start at the point (step) we are interested in?
  – Start at the learner’s last estimated scale location?
  – Start somewhere else?
Starting the Snapshot Assessment

- **Previous scale score available?**
  - **Yes**: Start at previous score
  - **No**:
    - **Step level specified?**
      - **Yes**: Start at step level
      - **No**: Start in middle of bank
A Snapshot Assessment

Step 2

9 September 2011 Symposium on Assessment and Learner Outcomes
Item selection

• Item selection algorithm
  – Select items around learner’s current estimate?
  – Select items around step level of interest?

• Content control
  – Progressions/topics
  – Text types
  – Snapshot around 1/3 of length of full CAT
    • Approx 1/3 of CAT content requirements
A Snapshot Assessment

Step 2

9 September 2011
Symposium on Assessment and Learner Outcomes
Termination Criteria

• When to stop?
  – Stop when we can be reasonably sure we have got a classification
  – We will never be certain!

• How sure do we need to be?

• How sure can we be?

• Want a short assessment. How short should we allow the assessment to be?

• How long should we go on before giving up on getting a classification?
Being “reasonably sure”

• Options

  – Sequential probability ratio test (SPRT)
    • The lowest acceptable score for a learner to be classified above
    • The highest acceptable score for a learner to be classified below
    • Takes care of false positives and false negatives with flexible probabilities
    • Focus is around the classification line

  – Confidence interval test
    • Simpler to apply than the SPRT
    • Interest is around the learner’s estimate rather than the classification line
Confidence interval test

95% chance that the true score lies in this range

5% chance that the true score lies in this range
Confidence interval test

95% chance that the true score lies in this range

5% chance that the true score lies in this range
A Snapshot Assessment

Step 2

9 September 2011

Symposium on Assessment and Learner Outcomes
How well does it work?

Correlation: 0.917
68.1% within 1 s.e
How well does it work?

Correlation: 0.917
68.1% within 1 s.e
How well does it work?

<table>
<thead>
<tr>
<th>Snapshot Assessment</th>
<th>Classified above</th>
<th>Unclassified</th>
<th>Classified below</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classified above</td>
<td>8453 (71%)</td>
<td>402 (3.4%)</td>
<td>5 (&lt;0.1%)</td>
</tr>
<tr>
<td>Unclassified</td>
<td>528 (4.4%)</td>
<td>1651 (13.9%)</td>
<td>206 (1.7%)</td>
</tr>
<tr>
<td>Classified below</td>
<td>2 (&lt;0.1%)</td>
<td>120 (1.0%)</td>
<td>535 (4.5%)</td>
</tr>
</tbody>
</table>
References and Contacts

• References

• Contacts
  – Charles Darr, Senior Researcher, NZCER  Charles.Darr@nzcer.org
  – Hilary Ferral, Senior Statistician, NZCER  Hilary.Ferral@nzcer.org