

項目反応理論 (IRT) について

前川眞一

東京工業大学
大学院社会理工学研究科
2013.12.09 中教審

アウトライン

- 日本的テスト文化、試験文化
- 項目反応理論 IRT

日本的テスト文化

- 年に一度、同一問題での試験の斉一実施
- 新作問題のみでの試験の実施
(プリテスト無し)
- 試験問題の公開
- 大問形式の利用
- 問題作成とテスト編集の融合
- テストの専門家の不在・不利用
- 素点・配点の利用 (0点と満点)

世界標準(?)

日本的試験文化の逆

- 年に一度、同一問題での試験の斉一実施
(年に複数回、異なる問題を使って分散実施)
- 新作問題のみでの試験の実施 (プリテスト無し)
(統計的性質の判明している良問のみを出題)
- 試験問題の公開
(試験問題は基本的に非公開 難易度等が変わらないように)
- 大問形式の利用
(独立項目と大問の混合)
- 問題作成とテスト編集の融合
(素材提供者、項目編集者、心理測定専門家)
- テストの専門家の不在・不利用
(各テストに数人の Ph.D が専属)
- 素点・配点の利用 (0点と満点)
(素点ではなく尺度得点の利用)。

日本の大規模試験の現状

- 司法試験、公務員採用試験、保健師助産師看護師、大学入試センター試験等
 - 日本的試験文化
- 医師国家試験（厚生労働省）
 - プール制（項目プール・項目バンク）へ移行中だったが・・・
- 情報処理技術者試験（経産省・（独）情報処理推進機構情報処理技術者試験センター）
 - **CBT (IRT)**で実施中
- 医療系大学共用試験（医療系大学間共用試験実施評価機構）
 - **CBT (IRT)**で実施中 薬剤師、看護師、獣医も開発中
- 日本留学試験（日本国際教育協会 **AIEJ** → 日本学生支援機構 **JASSO**）
 - **IRT**で実施中 尺度得点の利用
- 日本語能力試験（国際交流基金・公益法人日本国際教育支援協会）
 - **IRT**で実施中
- **JETRO** ビジネス能力日本語試験（日本貿易振興機構 **JETRO** → 漢検）
 - **IRT**で実施中 尺度得点の利用
- 法科大学院適性試験（**2003-2010** 日弁連、大学入試センター）
 - 本邦初の**競争原理**の導入だったが・・・

- 古典的テスト理論
Classical Test Theory
- 項目反応理論（項目応答理論）
Item Response Theory

テスト得点の性質

$$X = T + E$$

観測されるテスト得点

真の得点

誤差

この部分を取り除きたい

この部分だけほしい

テストの信頼性 =
X 中の T の割合

テストの等化と尺度化

- 等化

複数のテストフォームの得点を
比較可能にすること

- 尺度化

得点に意味を持たせること。

等化のための条件

- 共通受験者の存在

学力が同じと考えられる集団、
もしくは同じ受験生が
複数の異なるテストフォームを
受験していること。

- 共通項目の存在

複数のテストフォームに
同じ項目が含まれる。

テストのデザイン

- センター試験

	項目群1	項目群2	項目群3	項目群4	項目群5	項目群6	平均点
年度 1-1	■						■
年度 1-2		■					■
年度 2-1			■				■
年度 2-2				■			■
年度 3-1					■		■
年度 3-2						■	■

テストのデザイン

● 共通項目あり

	項目群1	項目群2	項目群3	項目群4	項目群5	項目群6	項目群7	平均点
年度 1-1	■	■						■
年度 1-2		■	■					
年度 2-1			■	■				
年度 2-2				■	■			
年度 3-1					■	■		
年度 3-2						■	■	

テストのデザイン

- 共通受験者あり

	項目群1	項目群2	項目群3	項目群4	項目群5	平均点
年度 1-1	■					■
M1-1	■	■				
年度 1-2		■				
M1-2		■	■			
年度 2-1			■			
M2-1			■	■		
年度 2-2				■		
M2-2				■	■	

複数のフォーム

● 医学系共用試験

	項目 1	項目 2	項目 3	項目 4	項目 5	項目 6
個人 1						
個人 2						
個人 3						
個人 4						
個人 5						
個人 6						
個人 7						
個人 8						

日本でのIRT (テスト理論) の歴史

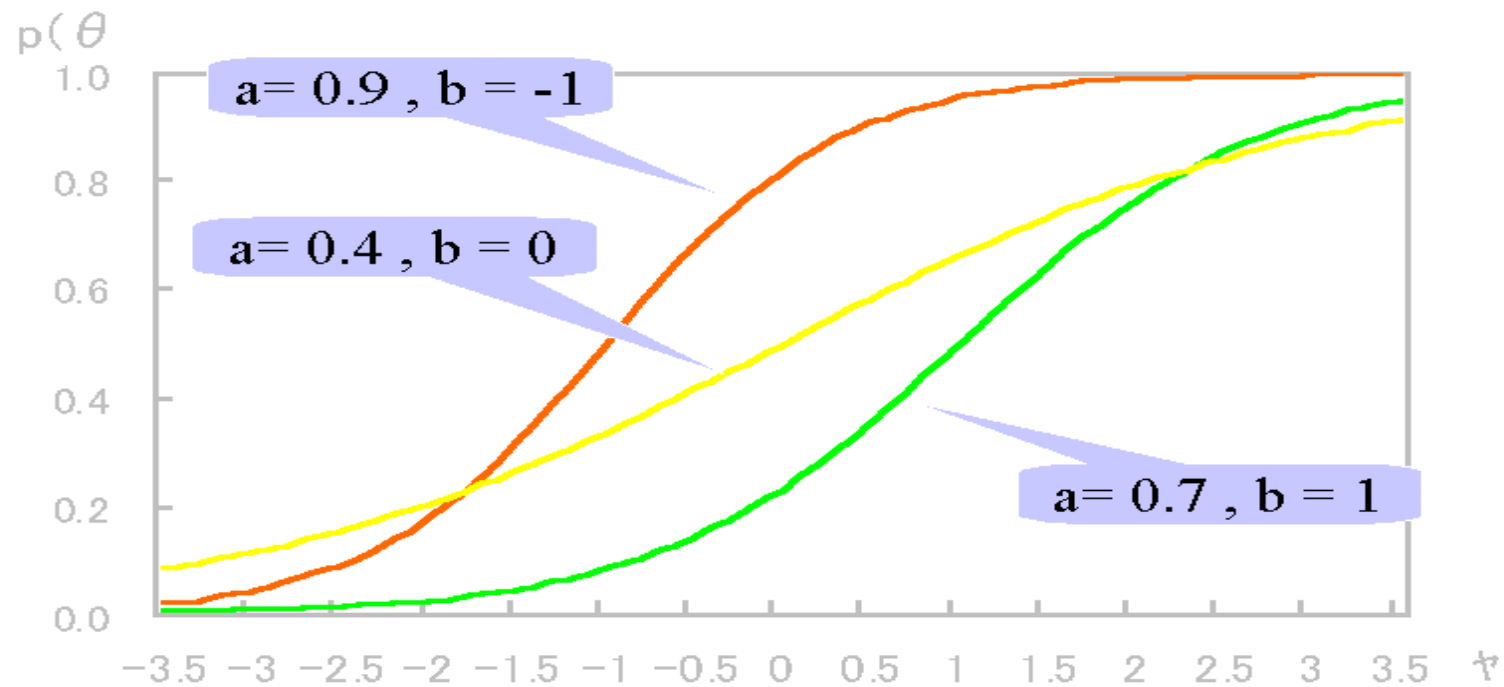
- 印東太郎・鮫島史子
- 池田央、芝祐順、柏木繁男、柳井晴夫
- 繁柵算男、渡部洋、石塚智一、村上隆
- 南風原朝和、野口裕之、村木英治、m
- 服部環、中村智靖、柴山直、豊田秀樹、植野真臣、熊谷龍一、荘島宏二郎、大久保智也・・・（他にも大勢おられます。）
- 民間企業等
 - 教育測定研究所、ベネッセ、リクルート、PROMETRIC、他

- テストに含まれる項目の統計的特性を調べる。 → 項目パラメタの推定
- 項目パラメタ
 - (集団に依存しない項目の統計的性質)
 - 項目困難度 (b) (正答率 (通過率) の逆)
 - 項目識別力 (a) (項目と合計点の相関)
- 項目特性曲線、テスト特性曲線

項目特性曲線

さまざまな項目特性曲線 (2PLM)

$$P_j(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))}$$

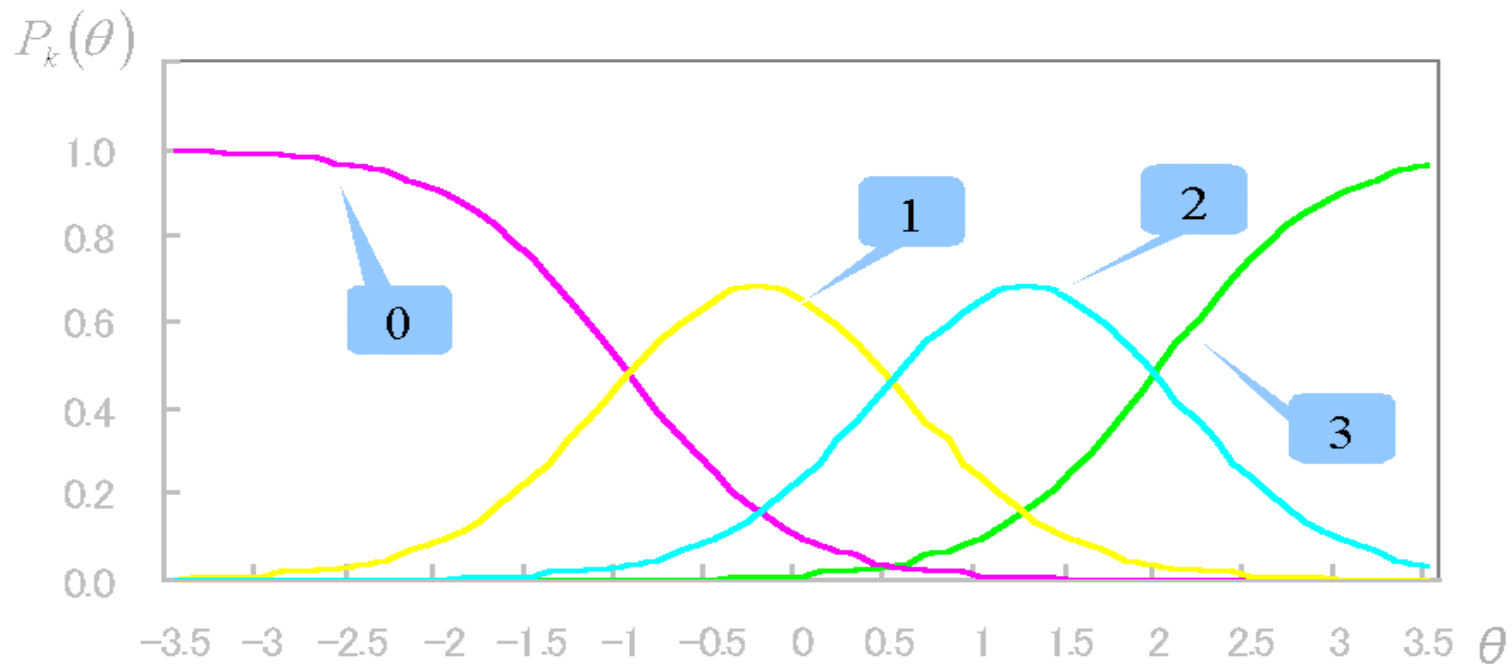


項目特性曲線（段階反応）

カテゴリ数4の場合の項目特性曲線

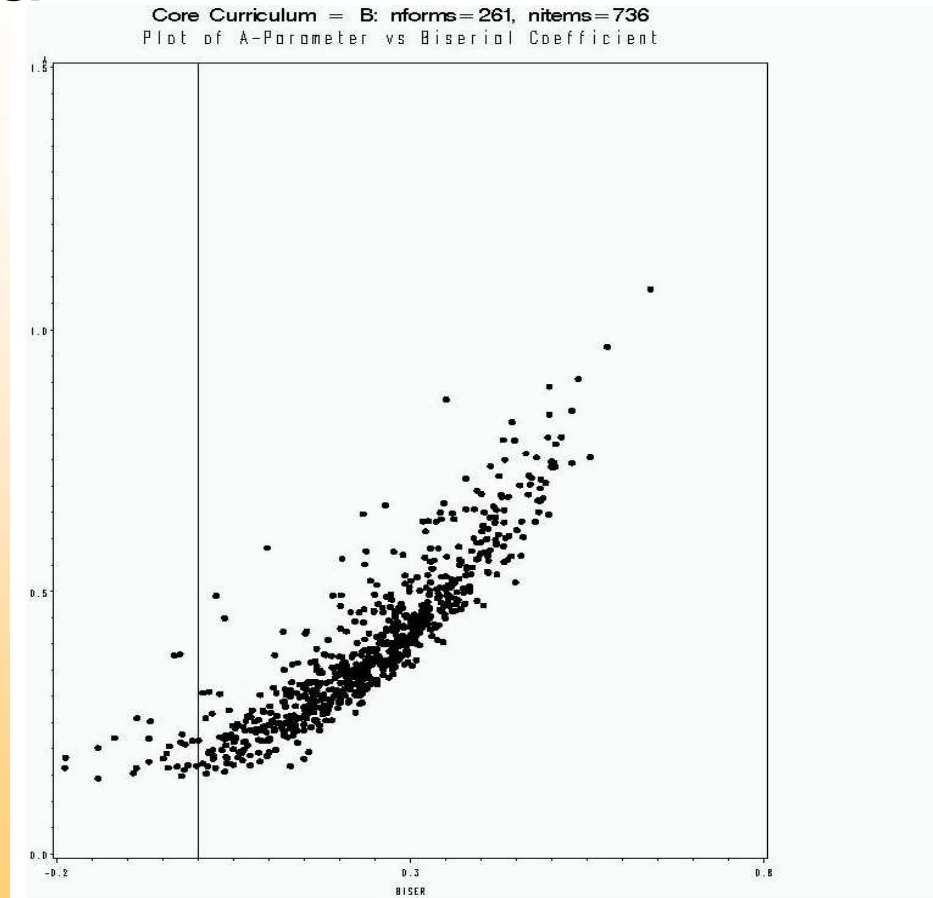
$$P_{jk}(\theta) = P_{jk}^+(\theta) - P_{jk+1}^+(\theta)$$

$$\text{ただし、} P_{jk}^+(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_{jk}))}$$



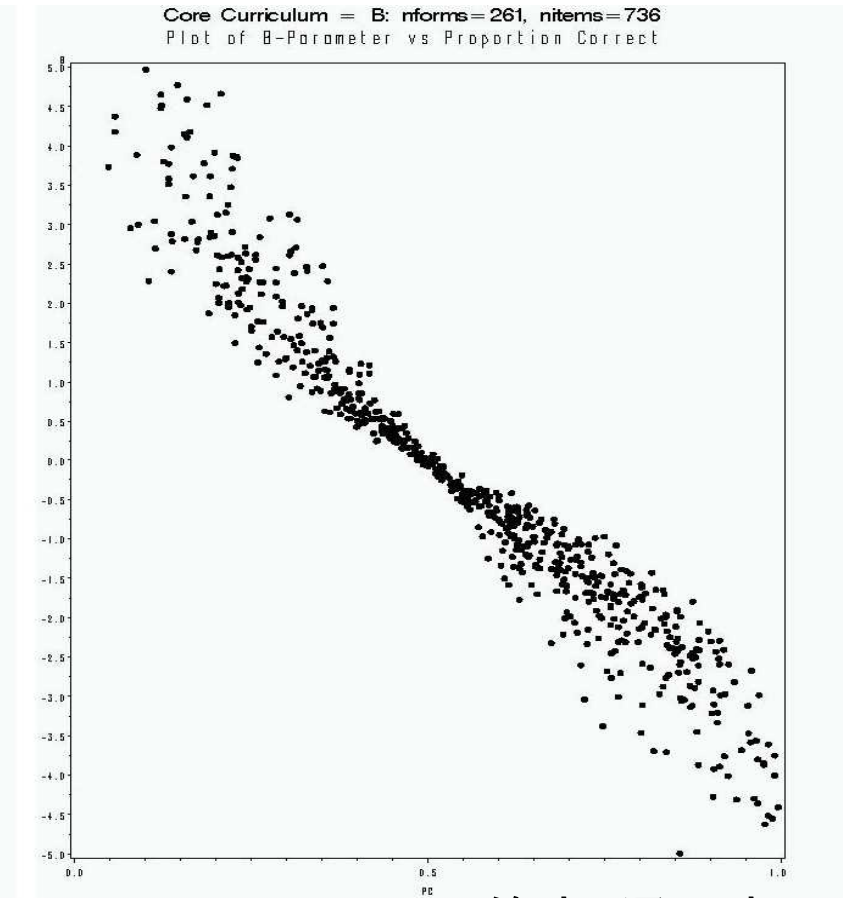
古典的項目統計量との関係

a



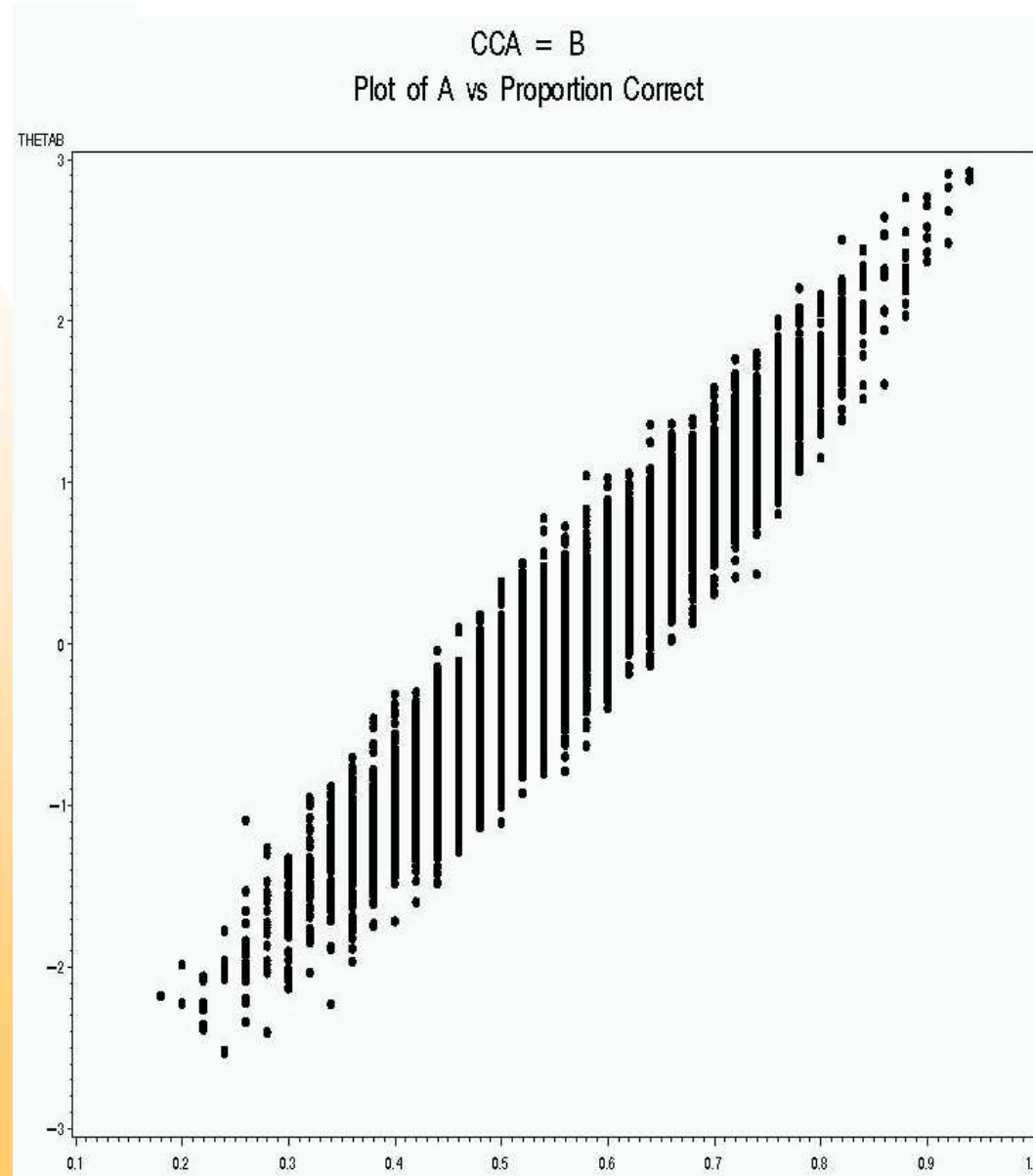
項目合計点相関

b



項目正答率(通過率)

正答数得点と θ の関係の例

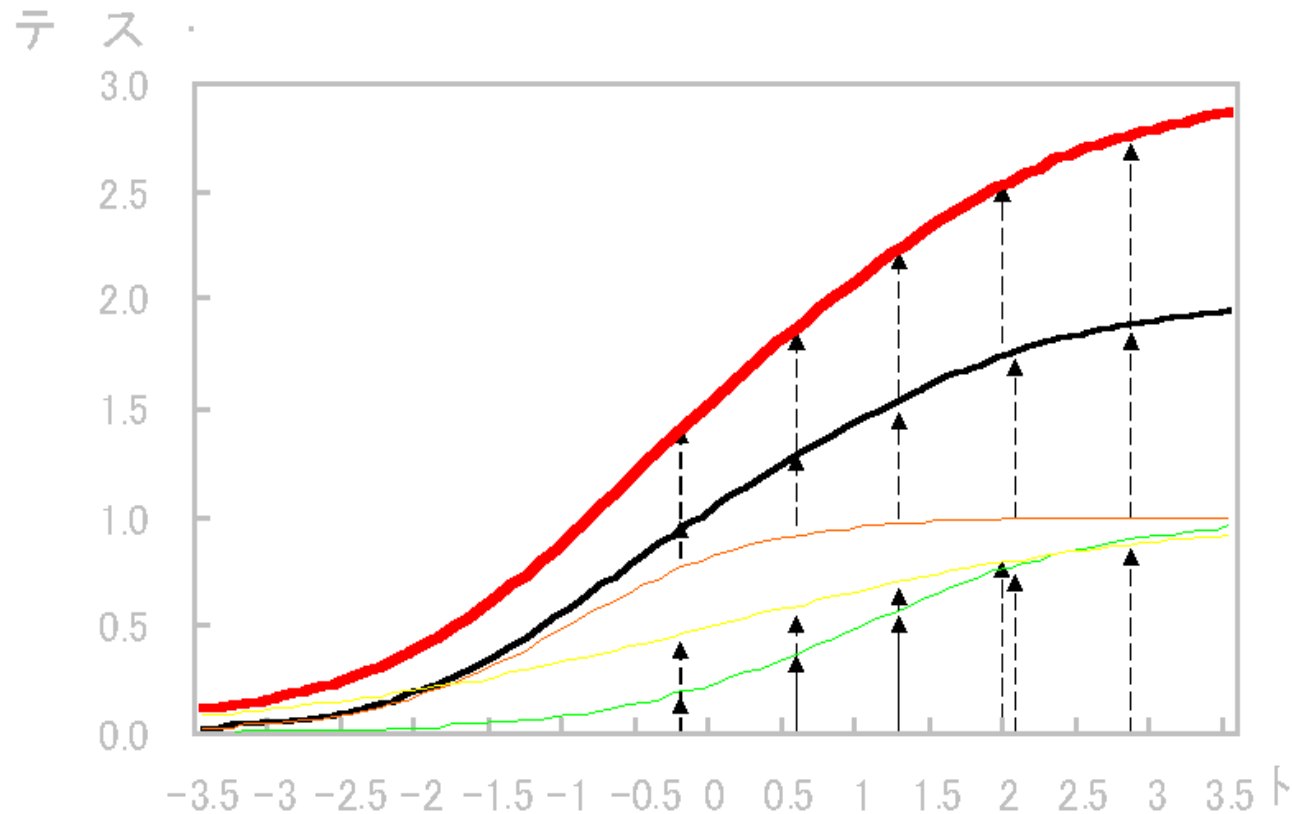
 θ 

テスト特性曲線

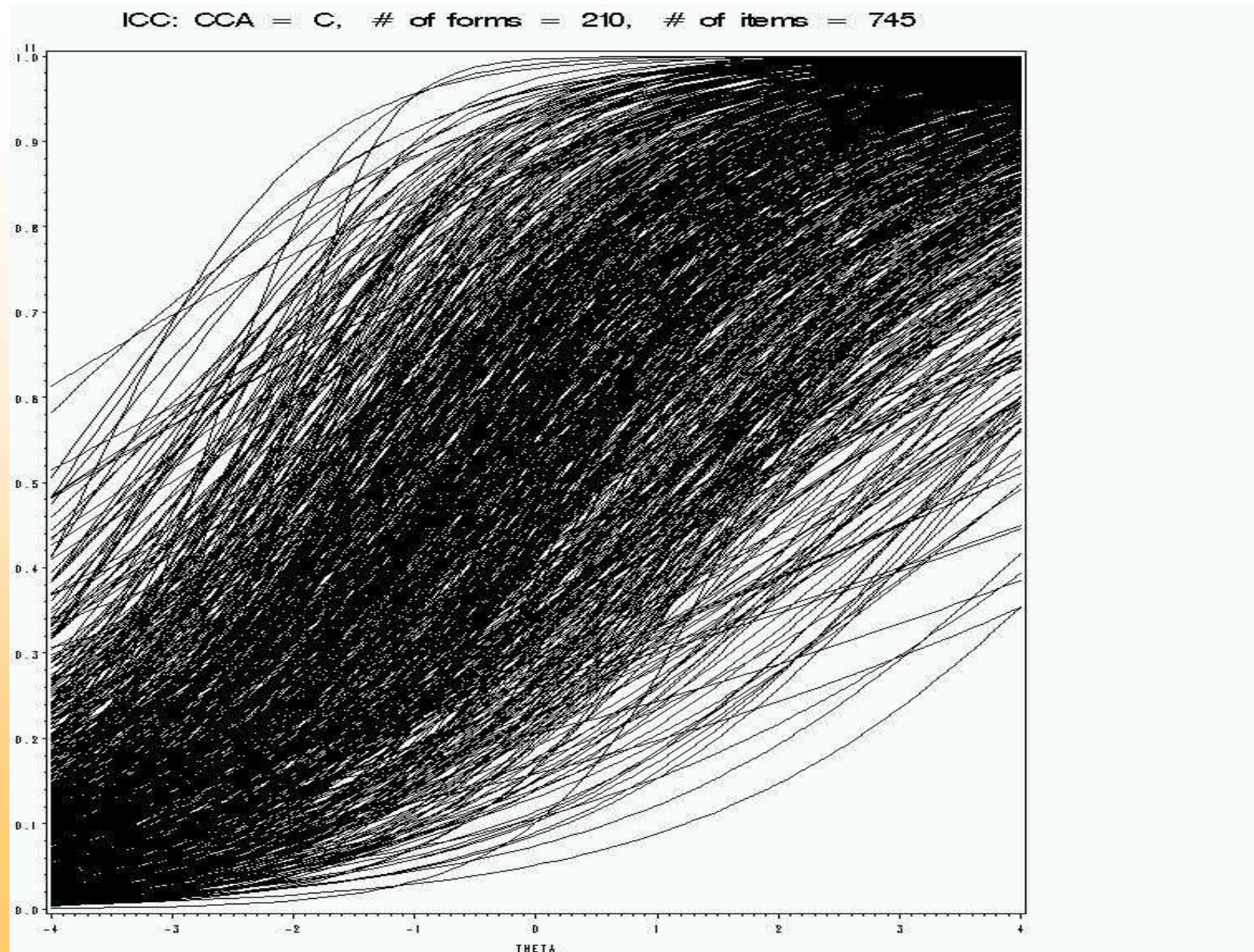
3項目からなるテストのテスト特性曲線

能力が θ の人が何点取るかを表す。

$$T(\theta) = \sum_{j=1}^p P_j(\theta)$$

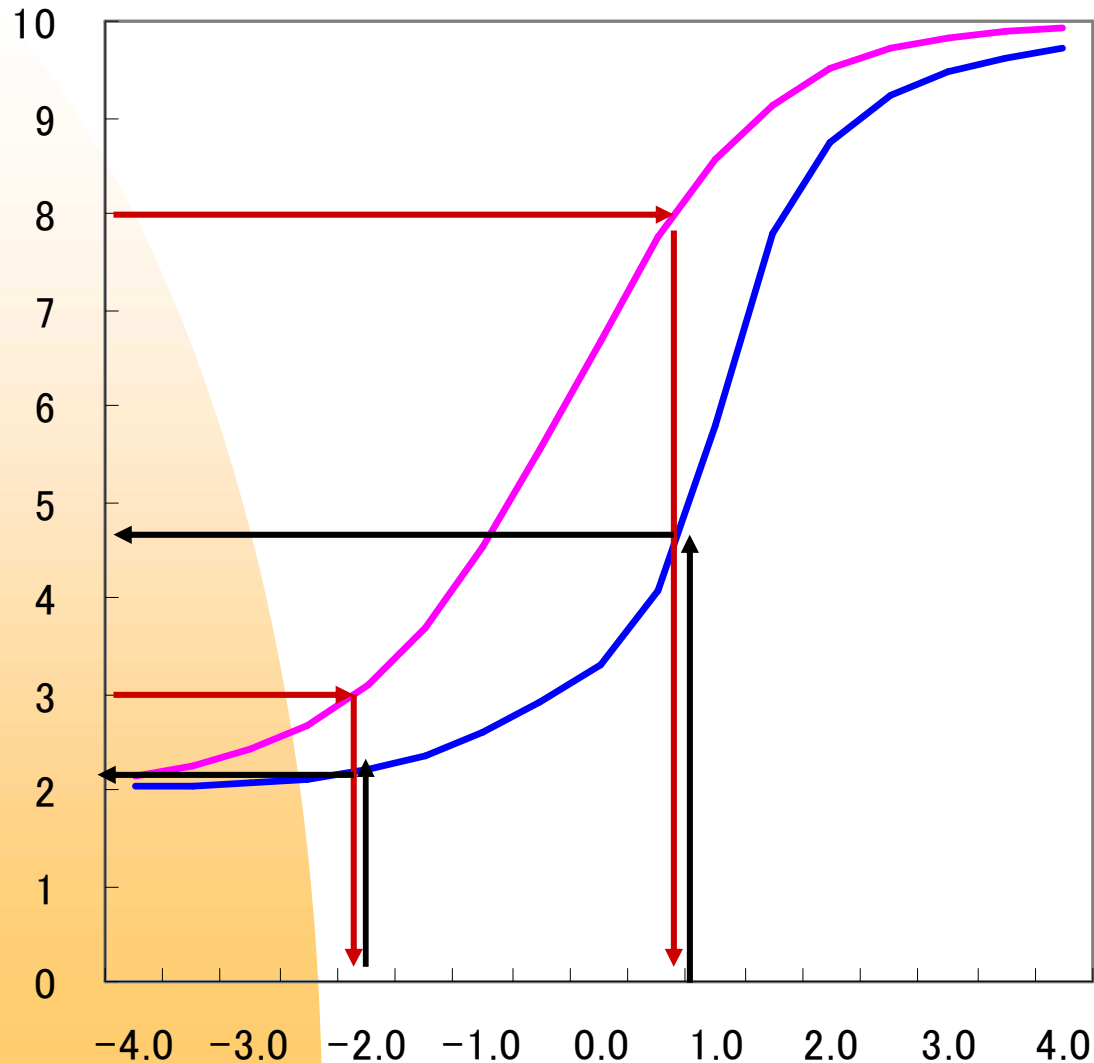


プール問題の項目特性曲線



フォーム間の比較（等化）

テスト得点



— TCCA

— TCCB

テストAの8点は
テストBの5点。
8点でも5点でも
 θ は1.0。

テストAの3点は
テストBの2点。
3点でも2点でも
 θ は-2.0。

θ

得点を比較可能にするには 1

23

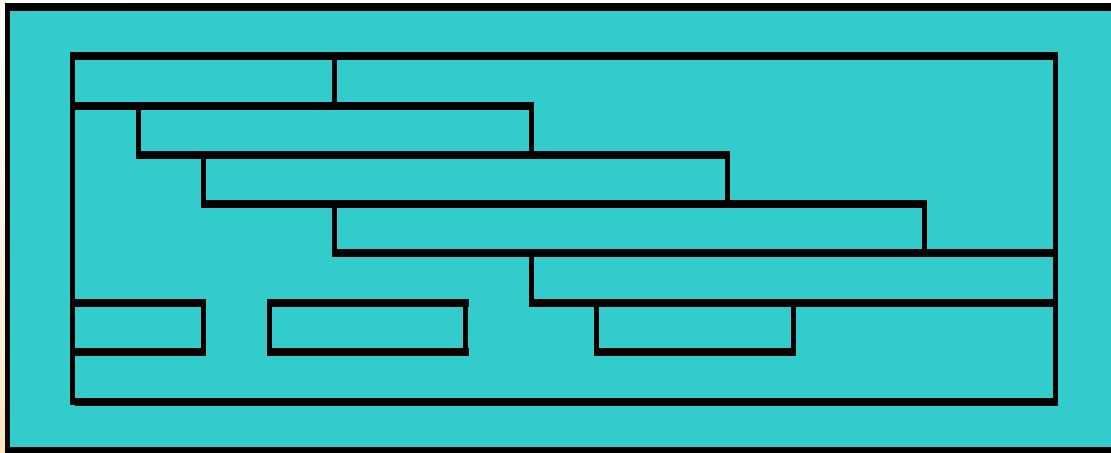
実施前に出題される項目の統計的特性
を調べておく → 項目のプリテスト

- 統計的特性（難易度・識別力等）が既知の項目のみを出題する
- 統計的特性（難易度・識別力等）が既知の項目をいくつか混ぜて出題する

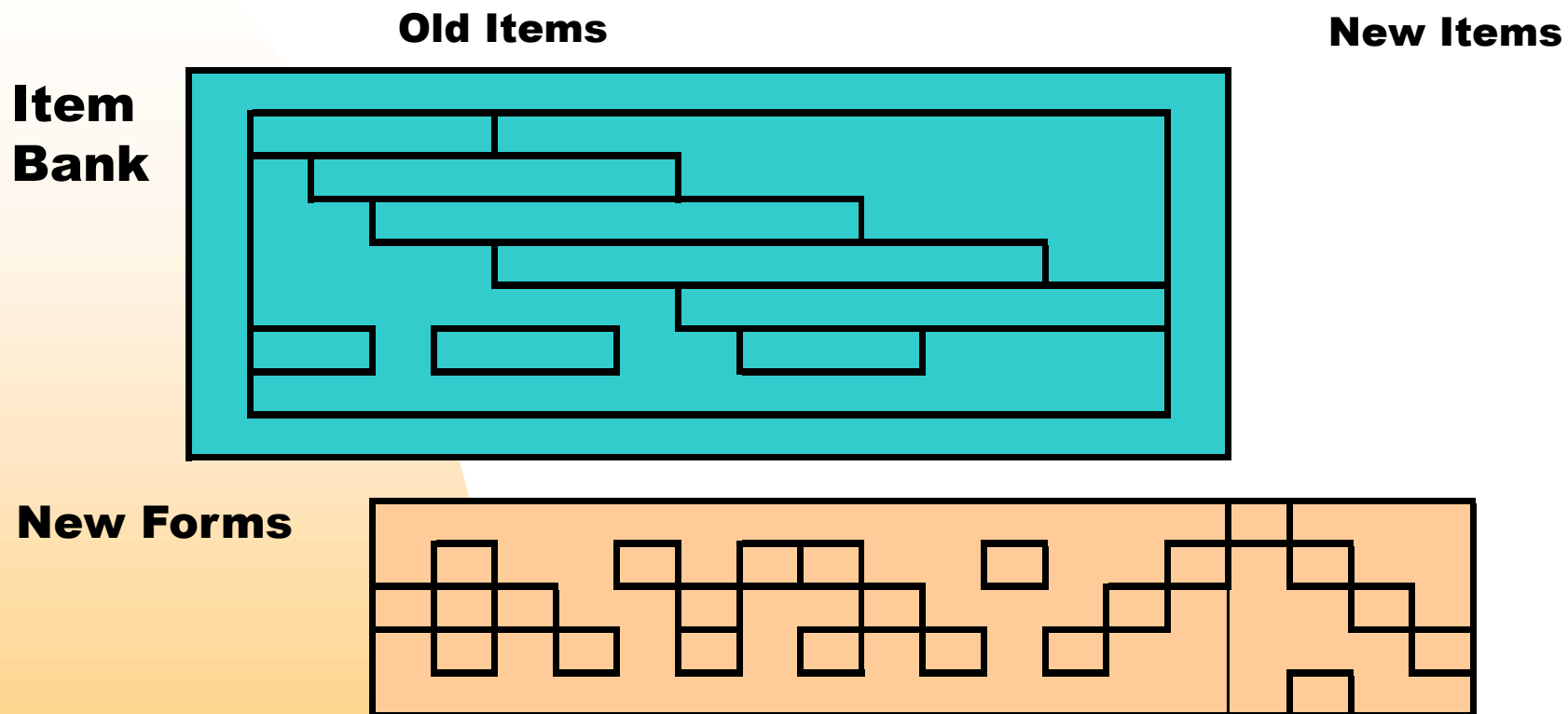
項目バンクの作成

form1
form2
form3
form4
form5
form6
form7
form8

item1 item2 item3 item4 item5 item6 item7 item8 item9 item1 item1 item1 item1 item1 item15



項目バンクの作成



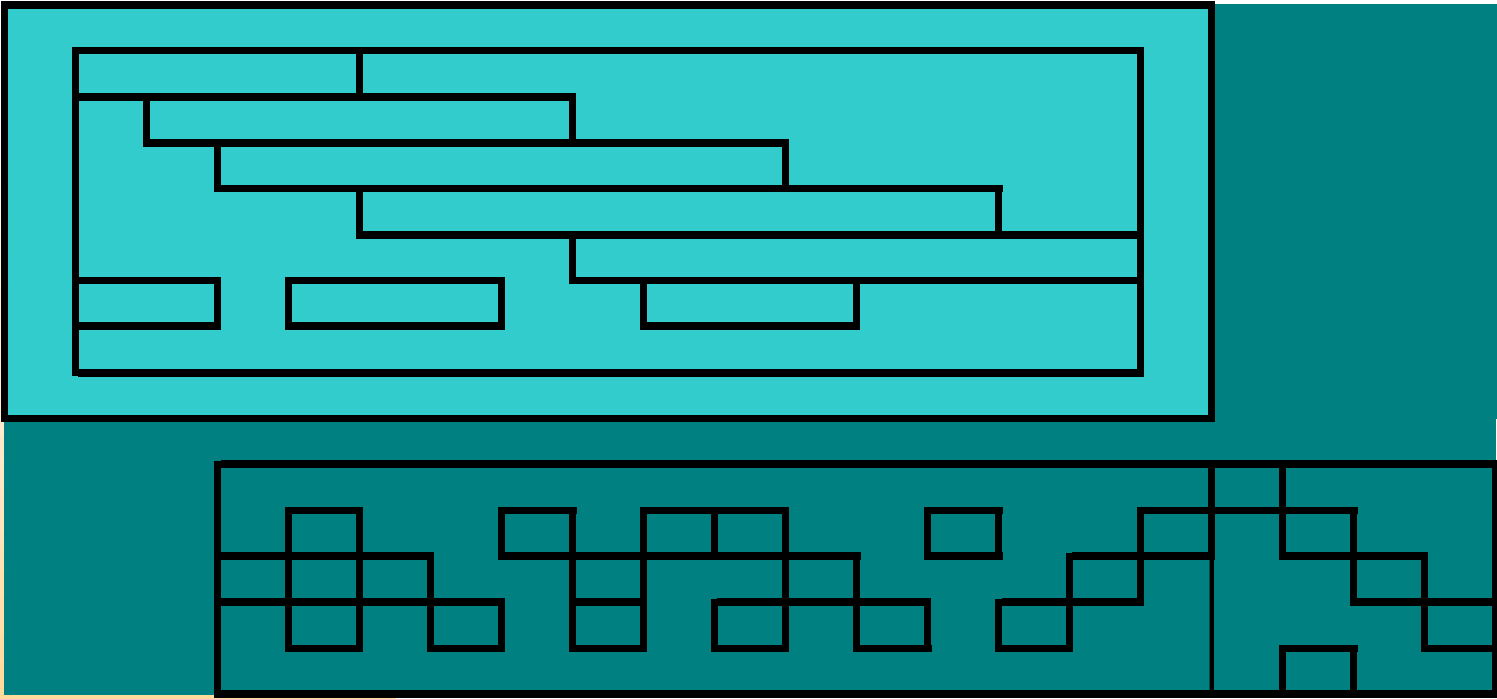
項目バンクの作成

新規
追加分

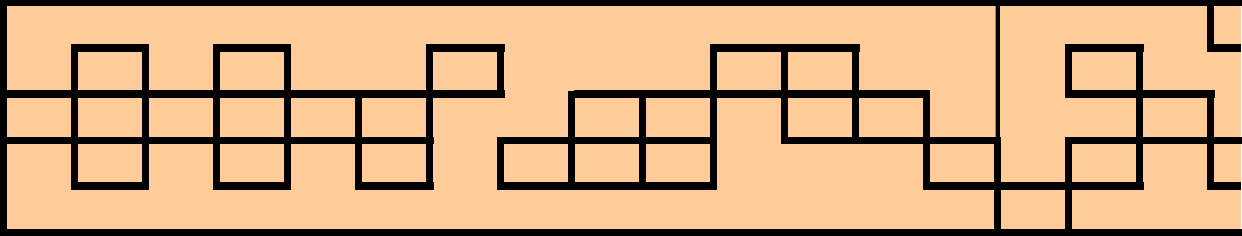
**New
Items**

Old Items

**Item
Bank**



New Forms



得点を比較可能にするには 2

27

実施後に、出題項目の統計的特性を調べる

- 試験の実施後、統計的特性を算出する
特定集団（共通受験者）によるモニター試験

モニター受験生に、新作問題と
統計的性質が既知の問題を、
同時に受験させる

テストのデザイン

● 共通受験者

	item1	item2	item3	item4	item5	項目群1	項目群2	項目群3	項目群4	平均点
年度 1-1						■				■
M1-1	■	■				■				
年度 1-2							■			
M1-2	■		■				■			
年度 2-1								■		
M2-1	■			■				■		
年度 2-2									■	
M-2-2	■				■				■	

	項目群1	項目群2	項目群3	項目群4	項目群5	項目群6	平均点
年度 1-1	■						■
年度 1-2		■					
年度 2-1			■				
年度 2-2				■			
年度 3-1					■		
年度 3-2						■	

たとえば

1. アンカーテストの作成ならびに
等化尺度化
2. 本試験の実施
3. 本試験と同時にモニター集団に本試験
とアンカーテストを実施
4. モニター試験のデータから、本試験を
アンカーテストに等化、尺度化
5. 尺度得点の通知
6. アンカーテストの更新、再尺度化

テストスタンダード（ガイドライン）

- テスト作成の規準
日本テスト学会ホームページより
ダウンロード可能

<http://www.jartest.jp/>

金子書房

ISBN-10: 4760828265

ISBN-13: 978-4760828265



- ご静聴有難うございました。