

全国的な学力調査のための 新しいタイプの調査方式について

東北大学大学院教育学研究科

柴山 直

(計量心理学・教育測定学)

平成23年度以降の全国的な学力調査の在り方 について(中間まとめ) 平成22年8月27日

- 検索 文科省 学力調査 中間まとめ 23
- http://www.mext.go.jp/b_menu/shingi/chousa/shotou/074/toushin/1297046.htm

- 7. 経年変化の分析等を重視した新しいタイプの調査方式の開発や、地方独自の調査との役割分担について

(1) 今後、調査目的の検討と関連させながら、経年変化の分析等を重視した新しいタイプの調査方式の開発を進める必要があると考えられる。

(2) また、全国的な学力調査に求められてきた調査目的の要素の一部を、今後は、地方独自の調査が担っていくシステムを構築することができないか、今後、検討を行う必要があると考えられる。

概 略

1. 現行制度の改善課題
 1. 暗黙の前提を見直す
 2. 制度設計のためのコンセプト
2. テストの役割分担
 1. Near is better
3. 全国的な学力調査のための最新のテスト技術
 1. 等化
 2. 対応づけ
 3. 項目反応理論
 4. 重複テスト分冊法

1. 現行制度の改善課題

最近
はB問題が
大切！

A問題の成績
が落ちる！

結果がすぐに
返ってこない！

国が作ってくれるので、
現場の指導力が落ち
る！

A問題：知識

B問題：活用

全国的な学力調査

PISA
ショック
・
学力低下
論争

経年比較ができない！
改善指標として使えない！

1.1 暗黙の前提を見直す

- 目的が**調査**なのか**指導**なのか
 - 「テストは競争の道具」であってはならない
 - 「テストは指導の道具」でなければならない
 - 「**テストは道具**」であって目的次第
- 全国一斉・同一冊子・同時実施
 - 道具としてのテストを支える「技術」観が固定
 - より柔軟な制度設計が必要
 - **最新のテスト技術を導入**

1.2 制度設計のためのコンセプト(例)

- 全国的な学力調査、自治体独自のテスト、教師作成テストの役割分担を明確にする (分ける)
- 必要であればそれらを相互に結びつける工夫をする (つなぐ)
- 経年比較を可能にする (続ける)

(注)「分ける・つなぐ・続ける」は京都府総合教育センターの松宮功氏の命名

2. テストの役割分担：“Near is better”

- 学校・教室： 教師作成（教師採点）テスト
 - － 児童・生徒に応じた
 - － 形成的評価ピンポイントの指導
- 自治体：独自テスト： アチーブメント・テスト
 - － 都道府県・政令指定都市ごと （悉皆）
 - － 地域・学校の実情に即した指導
- 国：全国的な学力調査： アセスメント
 - － 日本の子どもたちの学力をモニターする
 - － PISAやTIMSSと同様に目的はアセスメント（抽出）
 - － 経年比較 （等化）
 - － 学力の全体像 （重複テスト分冊法）

3. 全国的な学力調査のための 最新のテスト技術

- 等化 (Equating)
- 対応づけ (Linking)
- 重複テスト分冊法
(Item Matrix Sampling)
- 項目反応理論 (Item Response Theory)

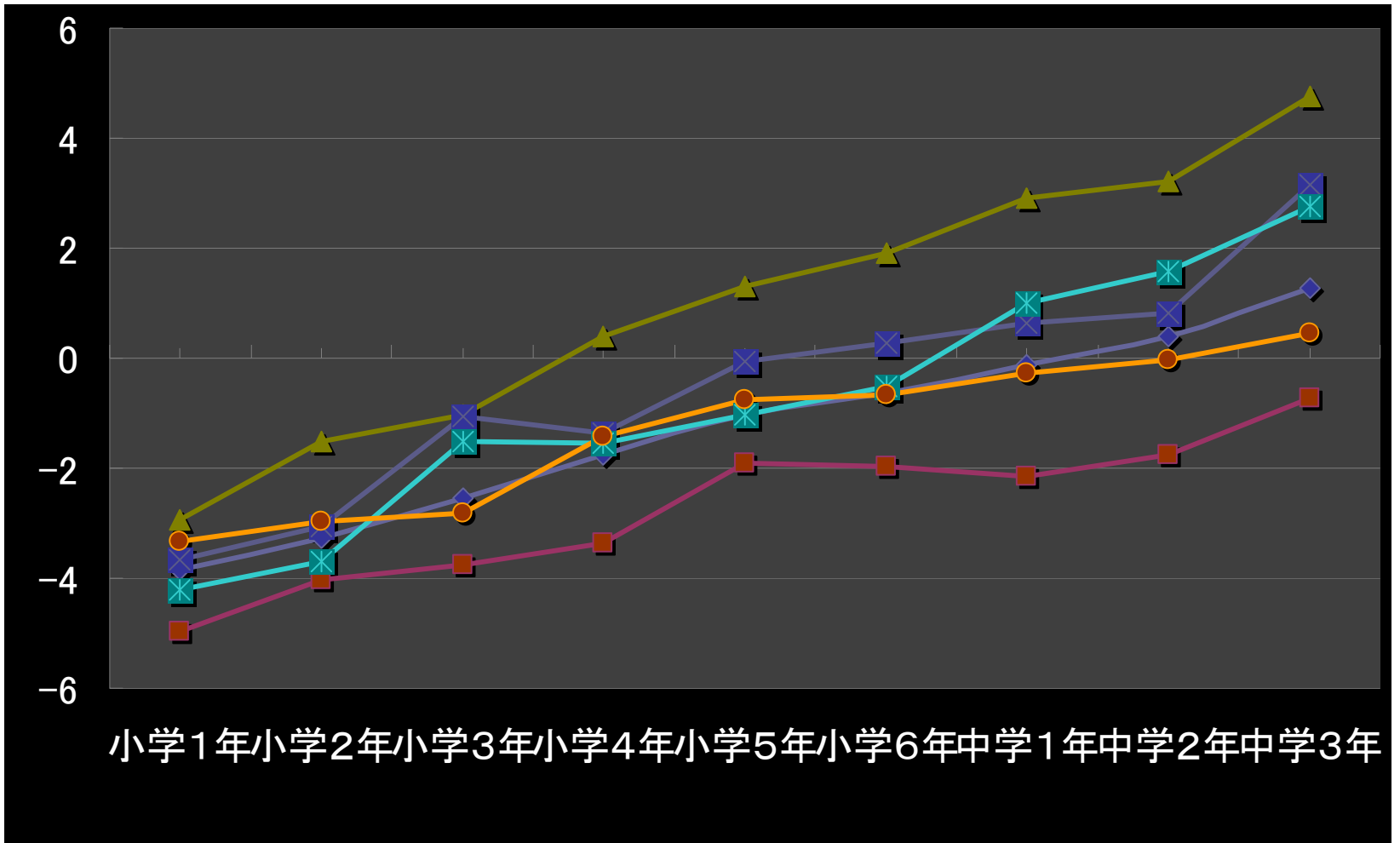
この他にサンプリング理論は必要

3.1 等化 (Equating)

～経年変化をモニターした例～

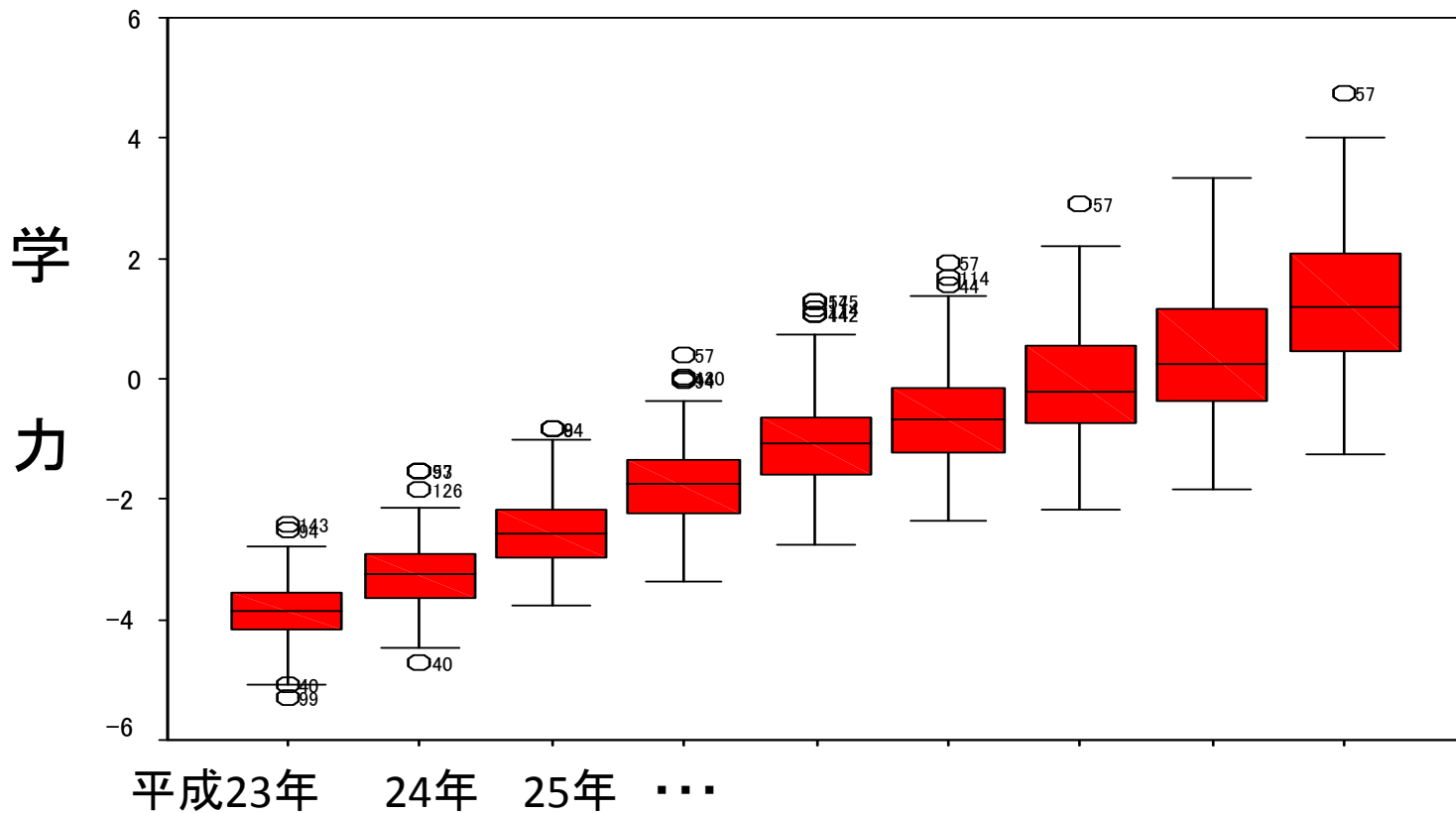
- 語彙理解力の発達に関する追跡的研究
- 東京都杉並区の公立小学校・中学校
- 入学時から卒業時まで追跡できた145名
- 項目間等化
 - － 学年で異なる項目の難易度
 - － 項目の難易度を共通尺度上に固定
 - － 理解力の変化を追跡できるようにした

語彙理解力の発達過程



学力の経年変化(イメージ)

年度で異なる項目の難易度
項目の難易度を共通尺度上に固定
学力の変化を追うことができるようになる



等化のためのデータ収集デザイン(基本)

1. 共通項目(テスト)デザイン: Non-Equivalent groups with Anchor Test Design

母集団	サンプル集団	テスト X	テスト Y	共通項目群
P	1	○		○
Q	2		○	○

2. 等価グループデザイン: Equivalent-Groups Design

母集団	サンプル集団	テスト X	テスト Y
P	1	○	
	2		○

3. 単一グループデザイン: The Single-Group Design

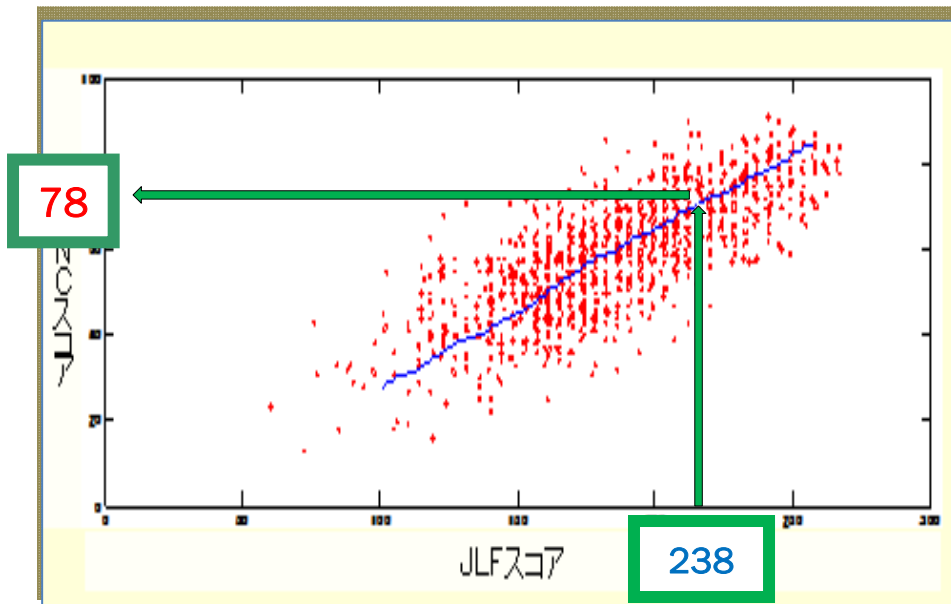
母集団	サンプル集団	テスト X	テスト Y
P	1	○	○

3.2 対応づけ(Linking)

～法科大学院適性試験～

- 法科大学院統一適性試験(JLF)
- 法科大学院適性試験(DNC)

等パーセンタイル
等化法の援用
(柴山・野口, 2004)

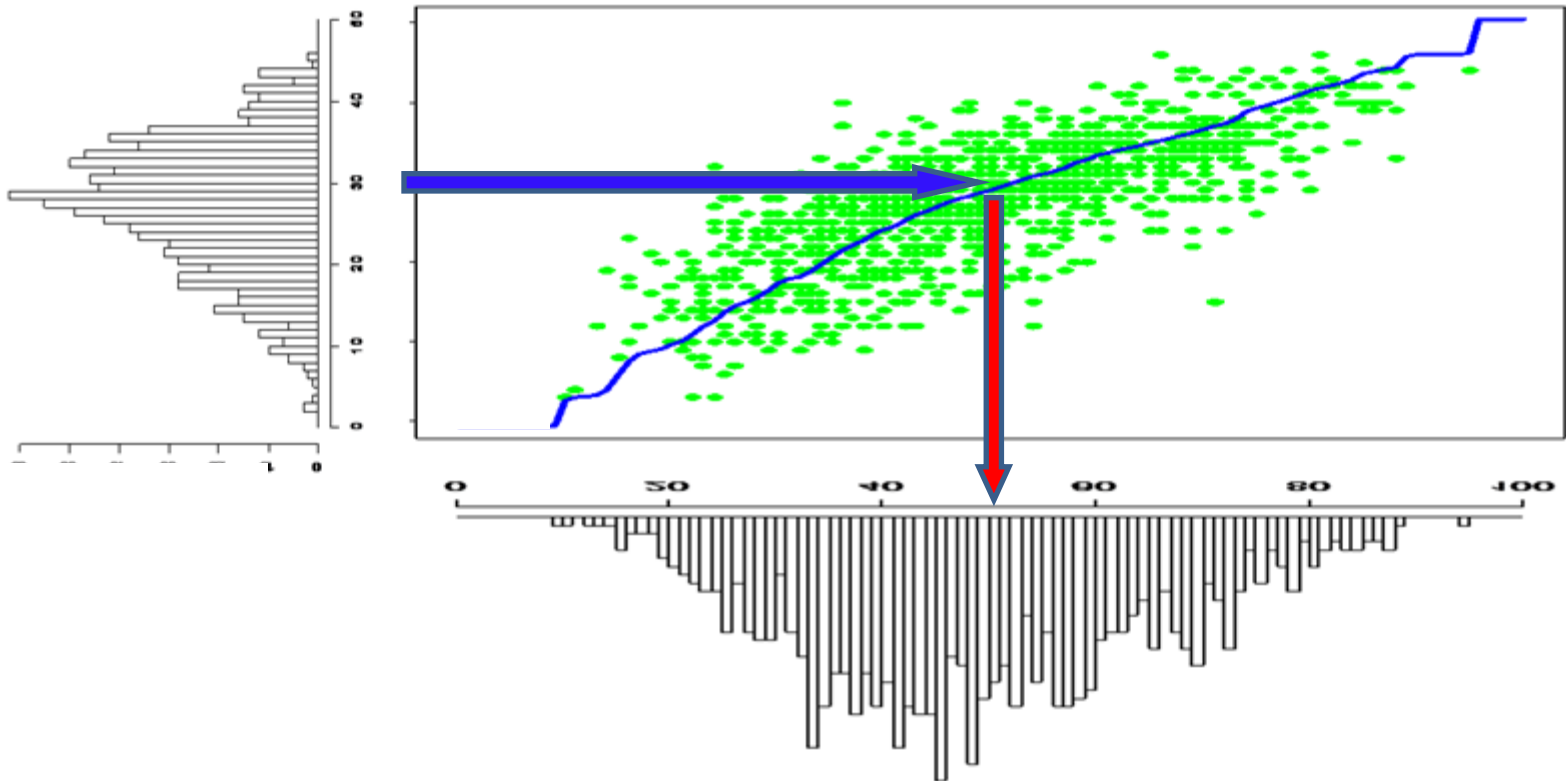


JLF スコア	DNC スコア	JLF スコア	DNC スコア
240	78.1	190	61.2
239	78.0	189	61.0
238	78.0	188	60.9
237	77.7	187	60.7
236	77.1	186	59.8
235	76.9	185	59.3
234	76.8	184	59.3

散布図と対応づけ曲線

全国的な学力調査と 自治体テストの対応づけ(仮想例)

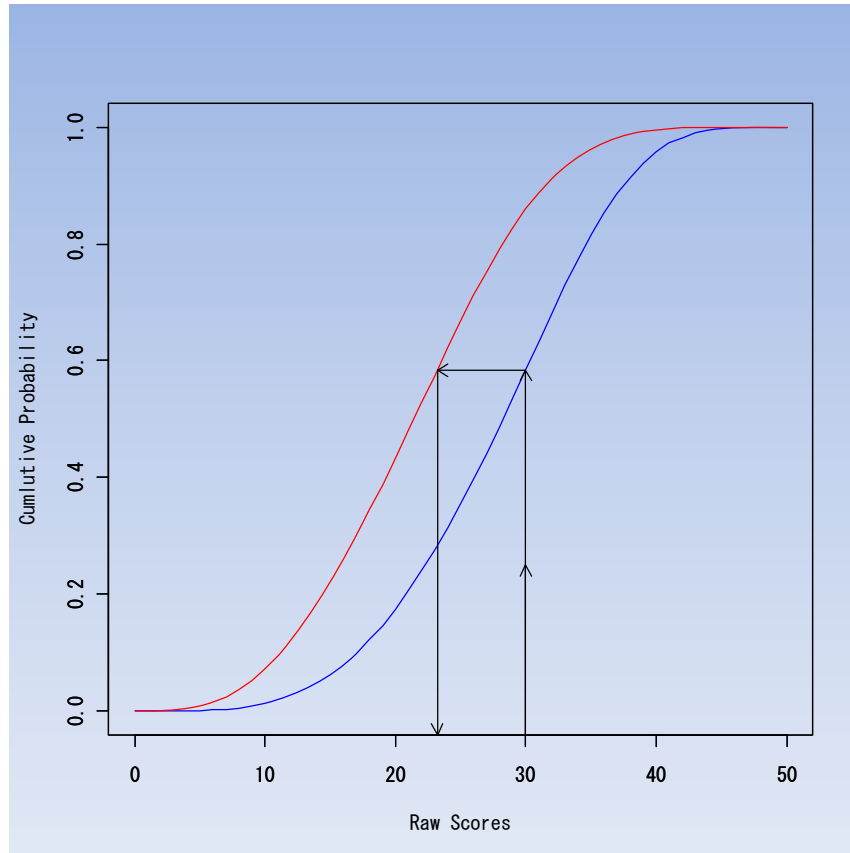
自治体テストの
得点分布



全国的な学力調査の
得点分布

等パーセンタイル等化法の原理を援用

累積確率（パーセンタイル）



正答数得点 X , Y

他にも

- ・線形等化法
- ・スムージングによる等化法
- ・IRTモデルにもとづく等化法

など...

3.3 項目反応理論 (Item Response Theory)

～IRTモデルを利用したテスト～

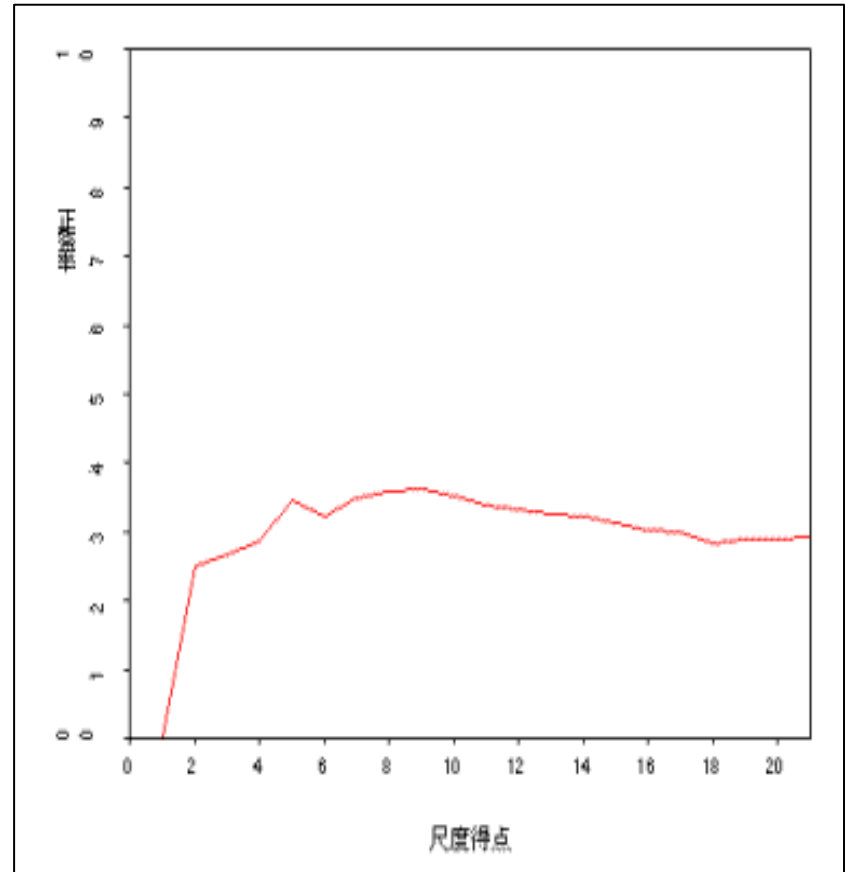
- TOEFL (Test of English as a Foreign Language)
- SAT (Scholastic Assessment Test:大学進学適性試験)
- LSAT (Law School Admission Test:法科大学院適性試験)
- GRE (Graduate Record Examination:大学院進学適性試験)
- OECD/PISA (生徒の学習到達度調査)
- TIMSS (国際数学・理科教育動向調査)
- NAEP (全米学力調査:全米教育進捗度調査)
- JLPT (日本語能力試験:日本国際教育支援協会)
- 医療系CBT (医療系大学間共用試験実施評価機構)

悪い問題の項目特性曲線

(個人差の識別ができていない)

- 問題 傍線部「夙に」の意味として最も適当なものを、次の①～⑤のうちから一つ選べ。

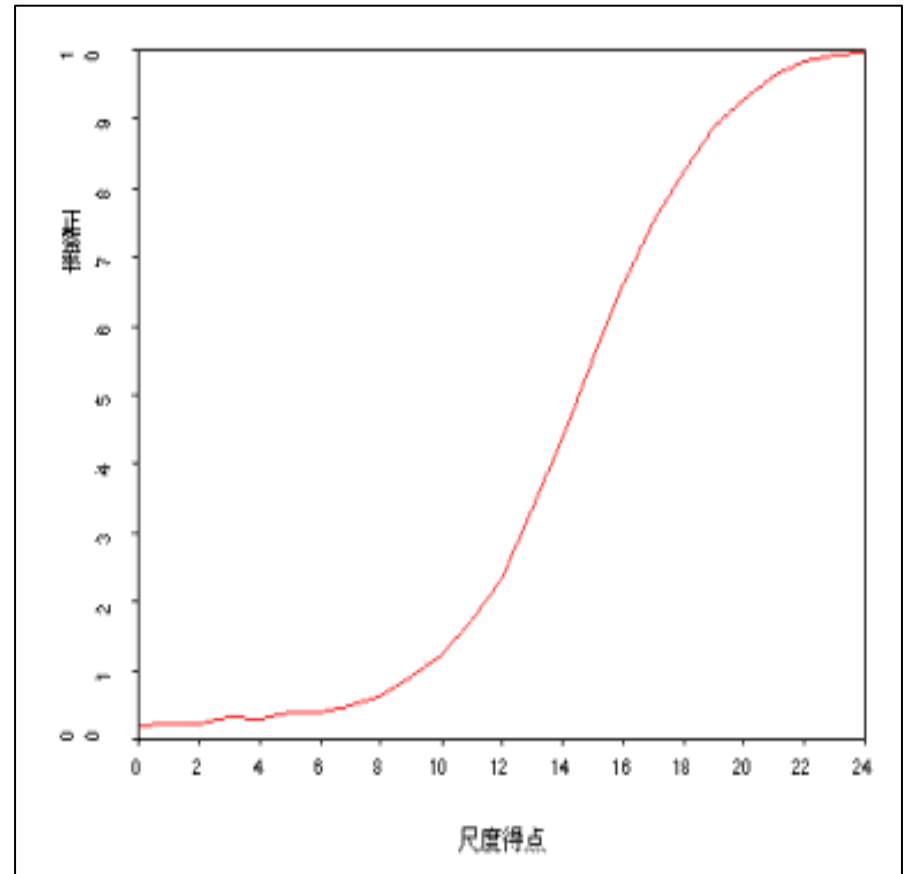
- ①うまれつき
- ②朝早くから
- ③心のそこから
- ④ずっと以前から
- ⑤十二分に



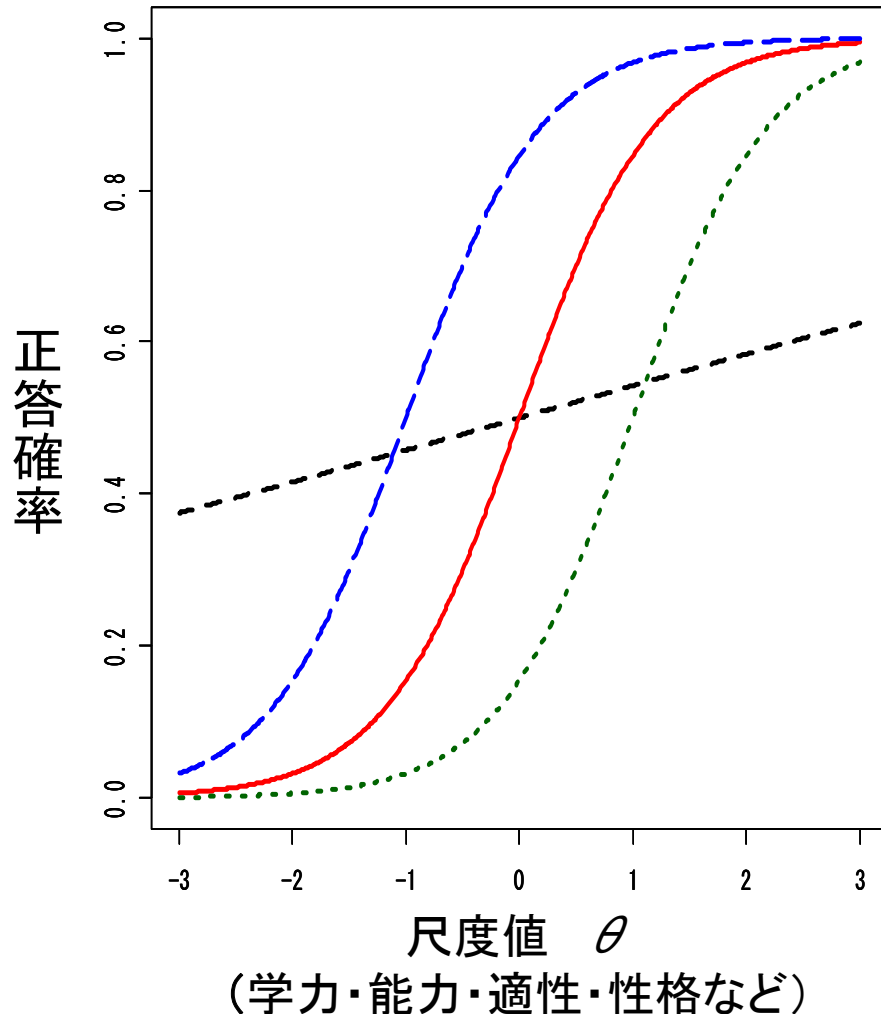
良い問題の項目特性曲線

(難易度:中)

- 問題 半径 $\sqrt{5}/2$ の円に内接する二等辺三角形ABCにおいて, $AB=AC=2$ とする. また, Aを通るこの円の直径をADとする. このとき, $\sin \angle BAC = \text{キ} / \text{ク}$ である.



項目特性曲線のモデル化

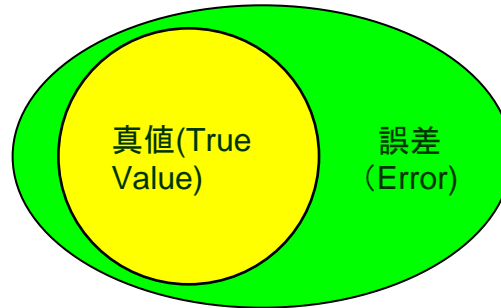


$$P(X_j = 1 | \theta) = \frac{1}{1 + \exp\{-1.7a_j(\theta - b_j)\}}$$

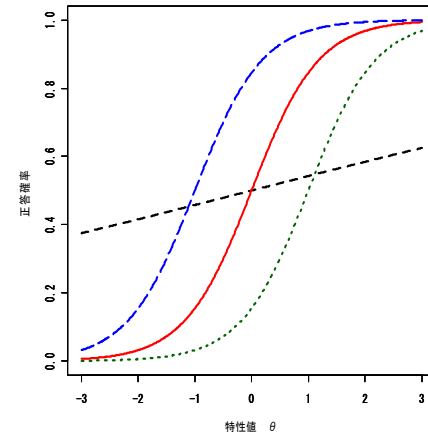
- (良い項目)
- 青い破線 困難度が低い (やさしい)
 - 赤い実線 困難度が中くらい (普通)
 - 緑い点線 困難度が高い (難しい)
- (悪い項目)
- 黒い破線 識別力が低い項目

古典的テスト理論と項目反応理論

$$\sum_{j=1}^m X_j = T + E$$

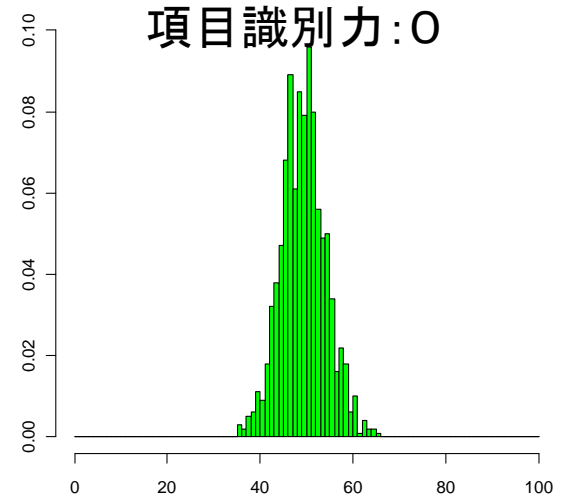
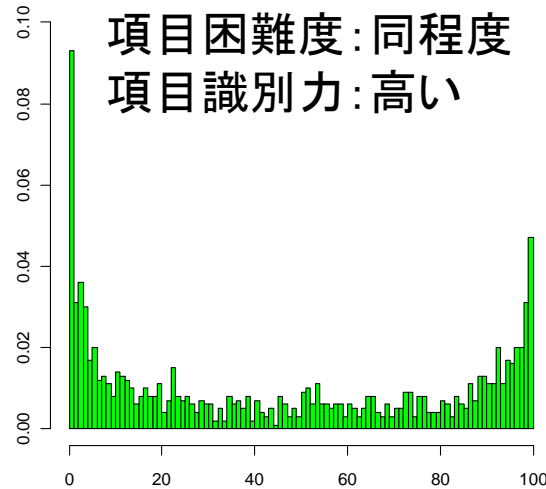
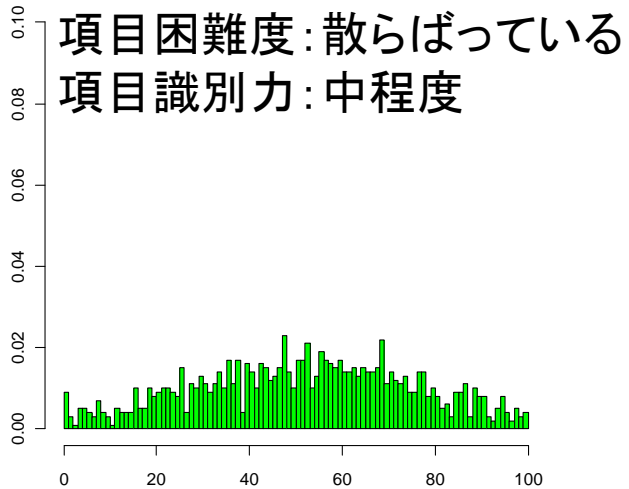
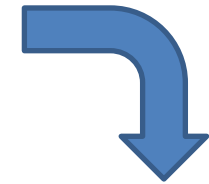
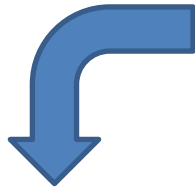
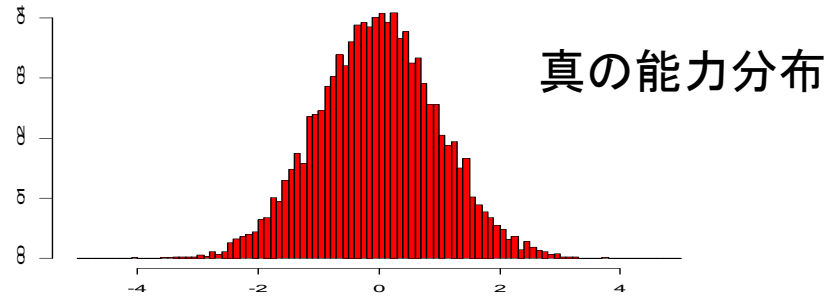


$$P(X_j = 1 | \theta) = \frac{1}{1 + \exp\{-1.7a_j(\theta - b_j)\}}$$



- 項目特性と受検者特性を分離
- 項目ごとに扱える
 - 柔軟なテスト設計が可能

得点の予測分布の例



テスト得点の予測分布

IRTモデルを利用したテストの特徴

- A) 実施は年に数度
- B) 実施手段によっては随時(CBT:IBTなど)
- C) 試作問題を実施時に組み込んでいる
- D) 問題は原則非公開
- E) 作成・品質管理・分析・報告に教科以外の専門家(Psychometrician・Statistician)も参加
- F) 素点ではなくて尺度得点
- G) 等化と併用することによって経年変化を追える

3.4 重複テスト分冊法 (Item Matrix Sampling)

～例：現行の学習指導要領の内容構成～

	A 数と式	B 図形	C 関数	D 資料 の活用	合計
第1学年	11	5	5	2	23
第2学年	6	5	4	2	17
合計	17	10	9	4	40

第2 各学年の目標及び内容(抜粋)

[第1学年]

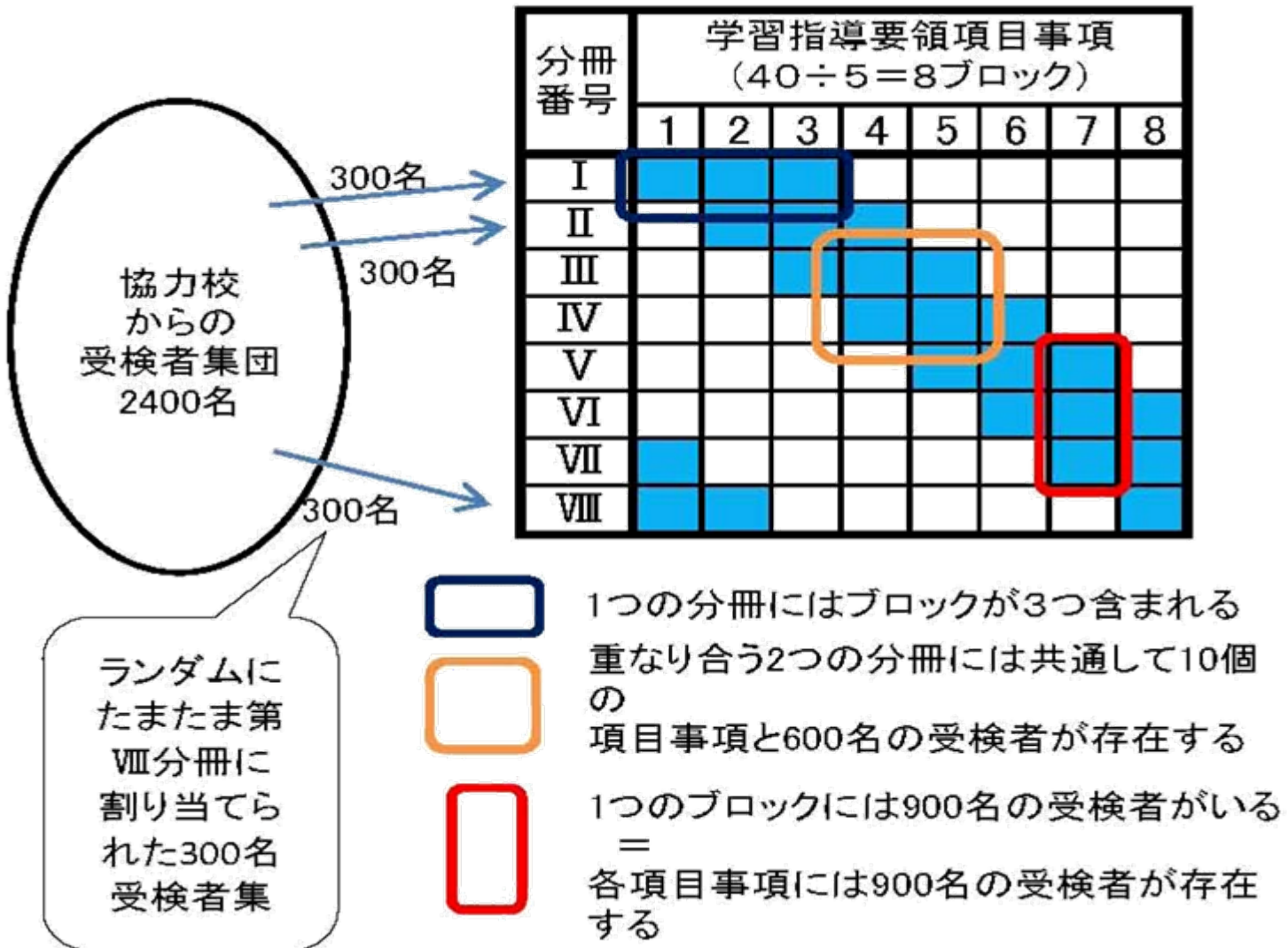
A 数と式

- (1) 正の数と負の数について具体的な場面での活動を通して理解し、その四則計算ができるようにする。
 - ア 負の数の必要性を知り、正の数と負の数の意味を理解すること。
 - イ 正の数と負の数の四則計算の意味を理解し、簡単な計算ができること。

.....

データ収集デザイン

御協力:新潟市教育委員会



まとめと今後検討すべき課題

- 学校・教室：教師作成（教師採点）テスト
- 自治体：独自テスト（アチーブメント・テスト）
↓ （対応づけ）



- 国：全国的な学力調査（アセスメント）
 - － 日本の子どもたちの学力をモニターする（抽出）
 - 経年比較 （等化：IRT）
 - 学力の全体像 （重複テスト分冊法：IRT）
 - － 新しい学力観・教育思潮への対応をどうするか
 - いまはPISA型「活用する力」
 - 時代に応じてまた変化していく
 - 全米教育進捗度調査(NEAP)のトレンド調査とメイン調査

ご静聴ありがとうございました