



経年変化分析調査との対応づけによる 本体調査の年度間比較の試み

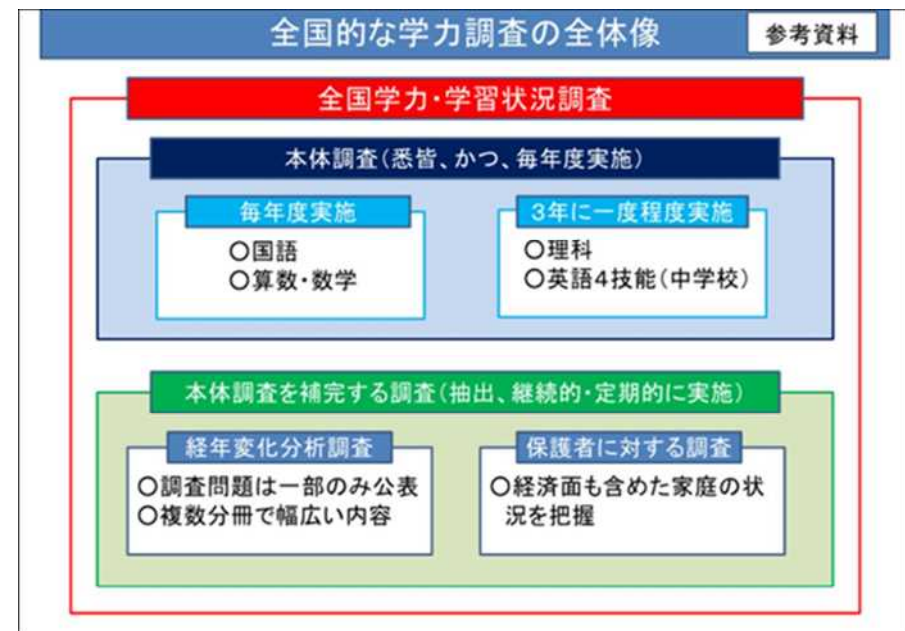
東北大学大学院教育学研究科
教育情報アセスメント講座

柴山 直

全国学力学習・状況調査における 本体調査と経年変化分析調査

- 本体調査
 - 全数調査・毎年実施
 - 問題公開
 - 国立教育政策研究所等より詳細な解説・報告書公開
 - 児童生徒の指導
 - 教育重点内容のメッセージ
 - 古典的テスト理論ベース
- 経年変化分析調査
 - 標本調査・定期的継続的
 - H25年度・H28年度実施済
 - 問題非公開
 - 幅広い領域をカバー
 - 学力の長期的変動の把握
 - 項目反応理論ベース

全国的な学力調査の全体像



「全国的な学力調査に関する専門家会議」文部科学省.(2013)

開発の目的

- A) 原理的には、年度内限りでの都道府県間の相対比較しかできない全国学力・学習状況調査の本体調査であっても、経年変化分析調査データを経由することで、都道府県単位ごとの年度間比較を可能とすること。

- B) 全国学力・学習状況調査の本来の目的の一つである政策立案 (Evidence-Based Policy Making: EBPM) のための総合的かつ客観的根拠を供給する技術的基盤を提供すること。

開発した3つの手法

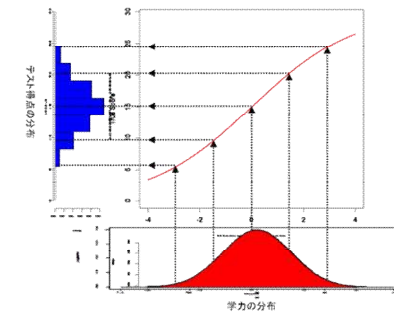
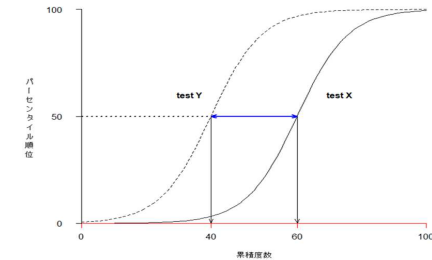
本体調査および経年変化分析調査のもつそれぞれの利点と、項目反応理論(Item Response Theory:IRT)モデルや 等化(equating)、対応づけ(linking)などの 学力調査の測定技術を組み合わせることで以下の3つの手法を開発した。

- 1) 平成25年度児童・生徒集団が仮に平成28年度本体調査を受検したと仮定した場合の復元得点分布 (IRT observed score distribution) による両年度間の学力分布の変化を捉える方法
- 2) 1) にもとづき、IRTモデルなどの複雑な方法を利用しなくても都道府県レベルで、上と同等のことが実現できる、平成25年度本体調査の得点を平成28年度本体調査規準の得点へ変換するための換算表(対応表: concordance table)を作成する方法
- 3) (追加提案手法の開発) 経年変化分析において、項目数が少ないため安定的な経年比較ができない 科目内領域ごとの年度間比較のために、推算値(Plausible Values:PVs)を利用する方法

開発手法に必須の3つのプログラミング

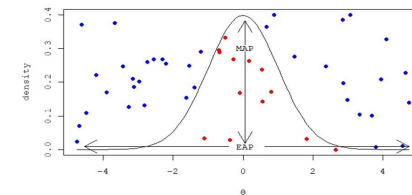
手法1と2の開発ために

- 等パーセンタイル等化法
 - 異なる2つテストの得点をパーセンタイル順位によって比較可能とする方法
- 復元得点分布 (IRT observed-score distribution) の生成方法
 - Recursion Formula (再帰式) を利用してH25年度集団のH28年度得点分布を復元 (推定) する方法



手法3の開発ために

- 推算値 (Plausible Values; PVs) を生成する方法
 - 項目が少ない場合にも母集団に関する統計量を偏りなく推定できる方法



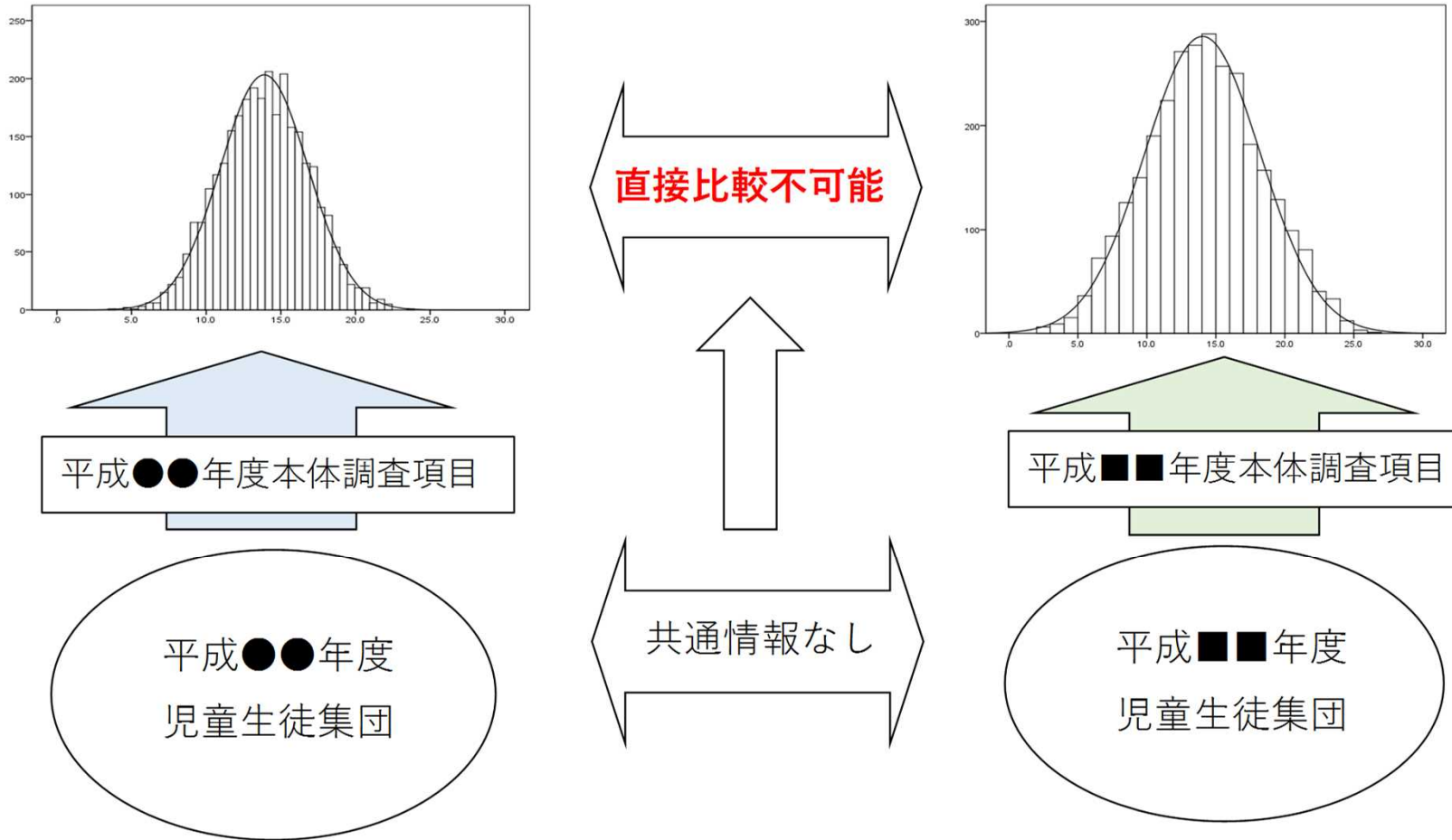
$$h(\theta|x) = \frac{f(x|\theta)g(\theta)}{\int f(x|\theta)g(\theta) d\theta}$$

研究組織

- **研究代表**
 - 柴山直(東北大学)
 - 研究統括・企画・プログラム設計・プログラム開発・報告書作成
- **研究協力**
 - 佐藤喜一(九州大学)
 - 理論的助言・専門的指導・プログラム作動検証
 - 熊谷龍一(東北大学)
 - EasyEstimationによる母数推定
 - 安永和央(九州大学)
 - 理論的助言
- **研究助手**
 - 澁谷拓巳(東北大学)
 - 板宮千尋(東北大学)
 - プログラム開発・データハンドリング・データ分析他
- **研究補助**
 - 江尻大亮(東北大学)
- **専門的助言**
 - 前川眞一(東京工業大学)
- **事務補佐**
 - 紙屋雅子(東北大学)

1 開発手続と分析対象データ

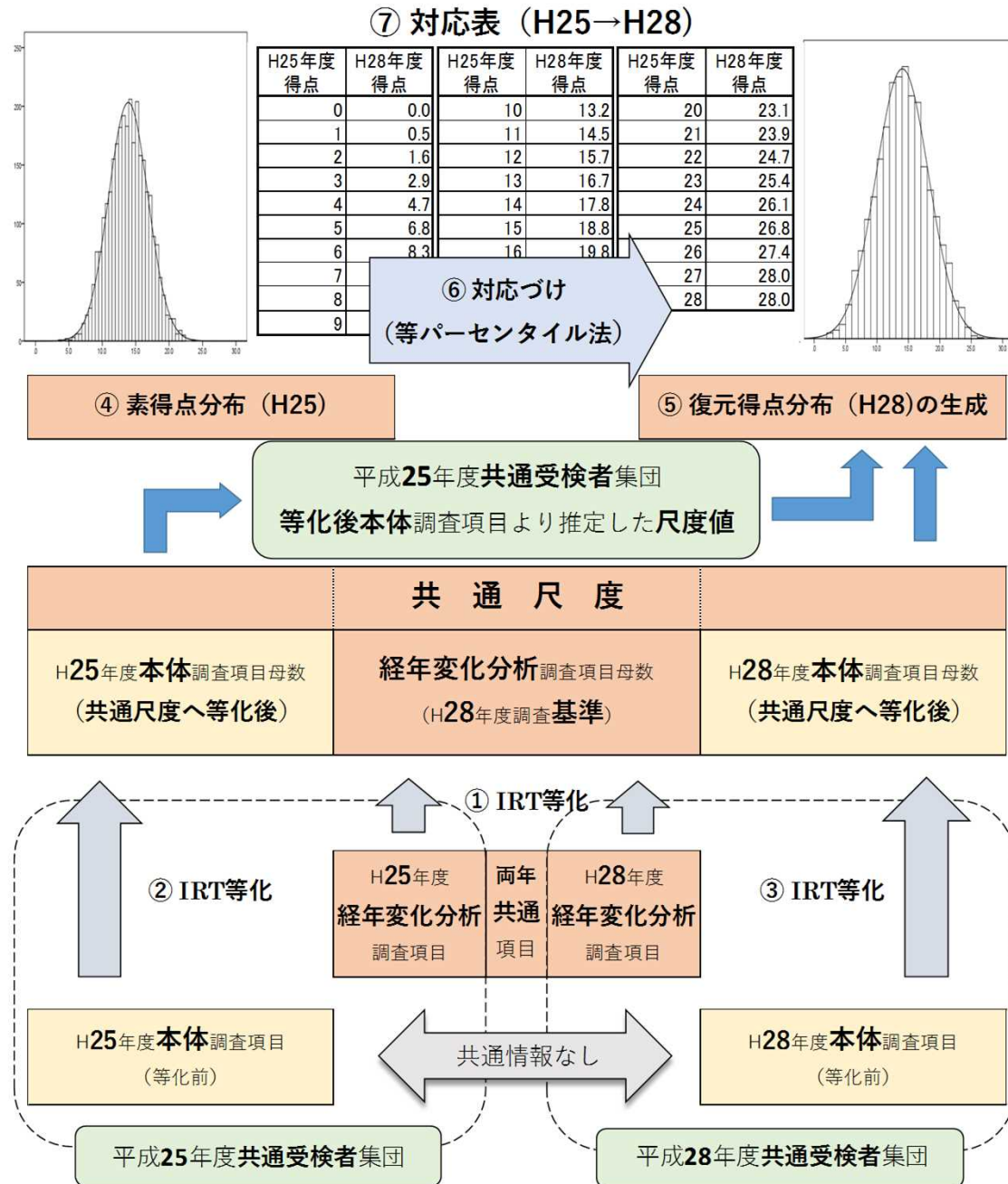
全国学力・学習状況調査の課題



研究 概念図

開発すべきプログラム

- ・等パーセンタイル等化 ⑥
- ・復元得点分布 ⑤
- ・推算値
(領域別経年変化分析)



データ収集デザイン(利用可能なデータ)

対象集団 年度	本体調査 (全数)		経年変化分析調査 (抽出)	
	平成25年度項目	平成28年度項目	H25年度項目	
			H28年度項目 (H25年度共通項目含む)	
平成 25 年度 集団		抽出	平成25年度 第1分冊 受検集団	
			平成25年度 第2分冊 受検集団	
平成 28 年度 集団		抽出		

抽出： 重複テ スト分 冊法： BIBD		1	2		3		4							1	分冊 番号 (受 検 集 団)
			1	2						3		4			
		1	2			3		4					3		
			1	3			4	2					4		
			4		1	2	3						5		
				4		1					3	2	6		
						1	3		4	2			7		
		3			4		1	2					8		
	2		3					1				4	9		
	4				3				1	2			10		
		4			2					1		3	11		
	3			4			2				1		12		
						2	3		4			1	13		
		1	2	3	4	5	6	7	8	9	10	11	12	13	
		項目セット													

分析対象とした調査データ

実施年度	調査種類	教科	学 年	
			小学6年生	中学3年生
平成25年度	本体調査 (全数)	国語	1,130,296	1,088,997
		算数/数学	1,130,730	1,089,359
	経年変化分析調査 (抽出)	国語	5,896	10,781
		算数/数学	5,881	11,605
平成28年度	本体調査 (全数)	国語	1,045,726	1,042,719
		算数/数学	1,046,363	1,042,929
	経年変化分析調査 (抽出)	国語	10,967	26,531
		算数/数学	10,753	25,942

全調査の項目数と項目母数の推定のためのサンプリング数

全調査項目数(共通尺度構成)

	小学校				中学校			
	国語		算数		国語		数学	
	本体	経年	本体	経年	本体	経年	本体	経年
H25年度	28	28	32	32	41	41	52	39
H28年度	25	39	29	52	42	52	51	65
共通項目数	0	24	0	28	0	38	0	36
非共通項目数	53	19	61	28	83	17	103	32

項目母数推定のためのサンプリング数

	小学校	中学校
平成25年度	11,777	22,386
平成28年度	21,720	52,473

2 本体調査における年度間比較の 実例

表記法

H25(raw) : H25の実データ分布そのもの グラフは黒点線

H28(raw) : H28の実データ分布そのもの グラフは黒実線

/ : (項目パラメタ群)/(尺度値群)

使う尺度値群がH25の場合はグラフは赤,

使う尺度値群がH28の場合はグラフは青

H25(H28) : H28年度基準の尺度(共通尺度上)で表現したH25項目パラメタまたは尺度値

例

H25(H28)/H25(H28) : H28年度基準の共通尺度上で表現されたH25年度項目パラメタ群に
H28年度基準の共通尺度上で表現されたH25年度尺度値群を代入して
復元したH25年度分布 グラフは赤

H28(H28)/H25(H28) : H28年度基準の共通尺度上で表現されたH28年度項目パラメタ群
(したがってH28年度項目パラメタそのもの)に
H28年度基準の共通尺度上で表現されたH25年度尺度値群を代入して
復元したH28年度分布 グラフは赤

H25(H25)/H25(H25) : H25年度項目パラメタと尺度値を使って復元したH25年度分布

H25(raw) vs H25(H25)/H25(H25) のグラフ:

: H25における実データ分布に対する復元得点分布の精度を見積もるための
グラフ 黒点線 対 赤

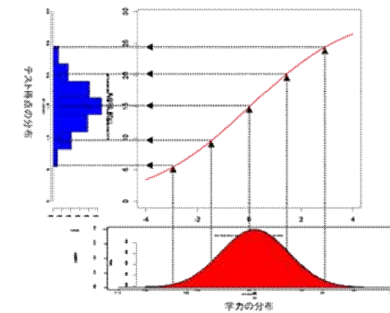
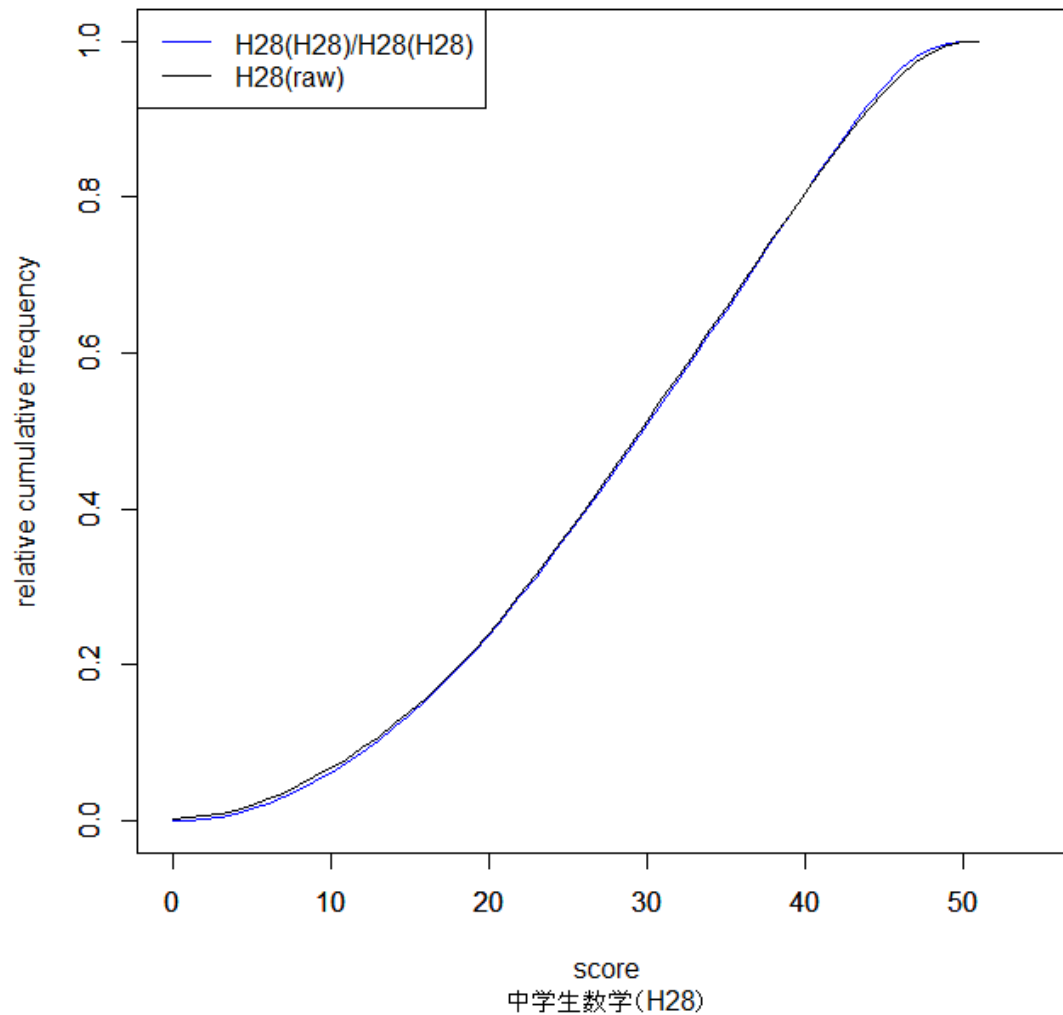
H28(raw) vs H28(H28)/H28(H28)のグラフ:

: 上と同様 黒実線 対 青

※ 尺度値 θ はEAP(expected a posteriori)推定法による推定値を利用している

実データ分布と復元得点分布の比較 による復元精度の確認

累積相対度数分布



H25年度集団がH28年度調査を受けたと仮定した場合の復元得点分布

復元得点分布 (H25→H28)

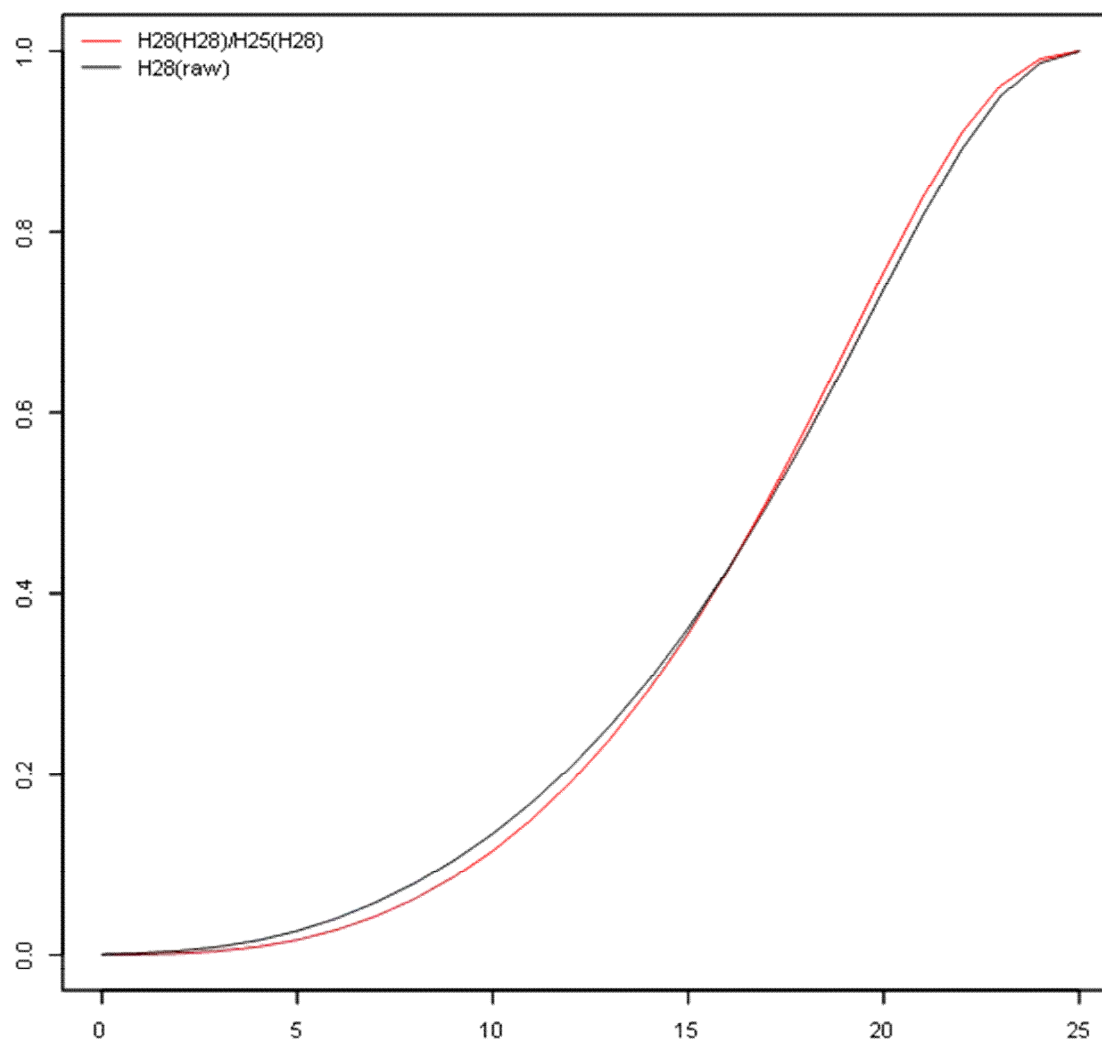
正答数	復元度数	正答数	復元度数
0	54.740	16	78092.790
1	406.620	17	85902.210
2	1360.700	18	92616.620
3	3039.640	19	97191.750
4	5458.710	20	98115.080
5	8627.880	21	93388.160
6	12503.410	22	80861.450
7	16975.350	23	59565.530
8	21935.340	24	32674.150
9	27333.500	25	9690.680
10	33182.710		
11	39527.830	合計	1130295.990
12	46406.600		
13	53817.010		
14	61691.990		
15	69875.540		

※予測値のため小数点以下第3位まで表示

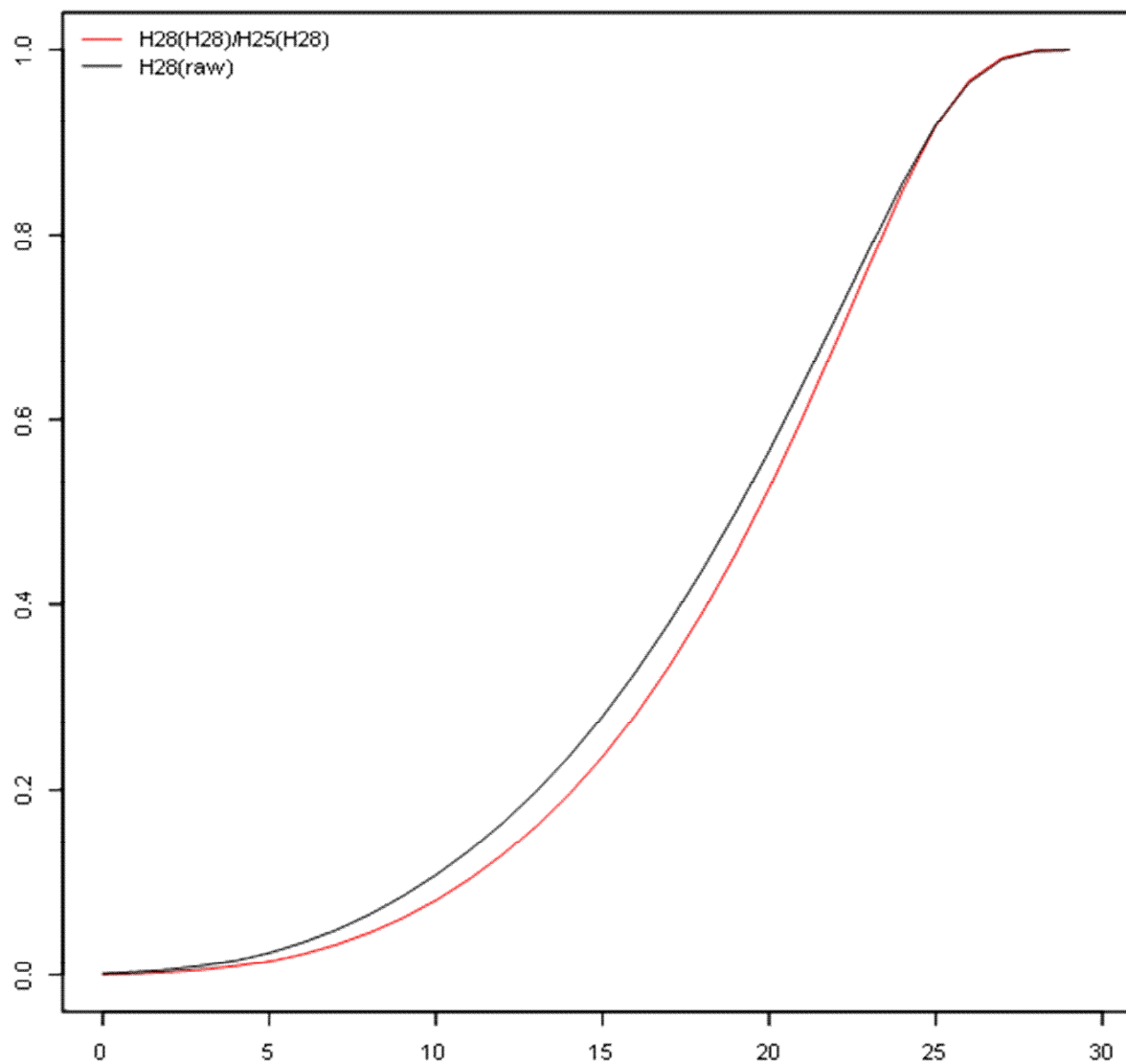
H28年度実データ分布

正答数	度数	正答数	度数
0	904	16	67,044
1	1,350	17	73,209
2	2,670	18	79,873
3	4,735	19	83,968
4	7,416	20	87,257
5	10,906	21	85,212
6	14,442	22	77,135
7	18,427	23	61,699
8	22,250	24	38,059
9	26,891	25	13,587
10	30,836		
11	36,043	合計	1045726
12	41,568		
13	47,160		
14	53,203		
15	59,882		

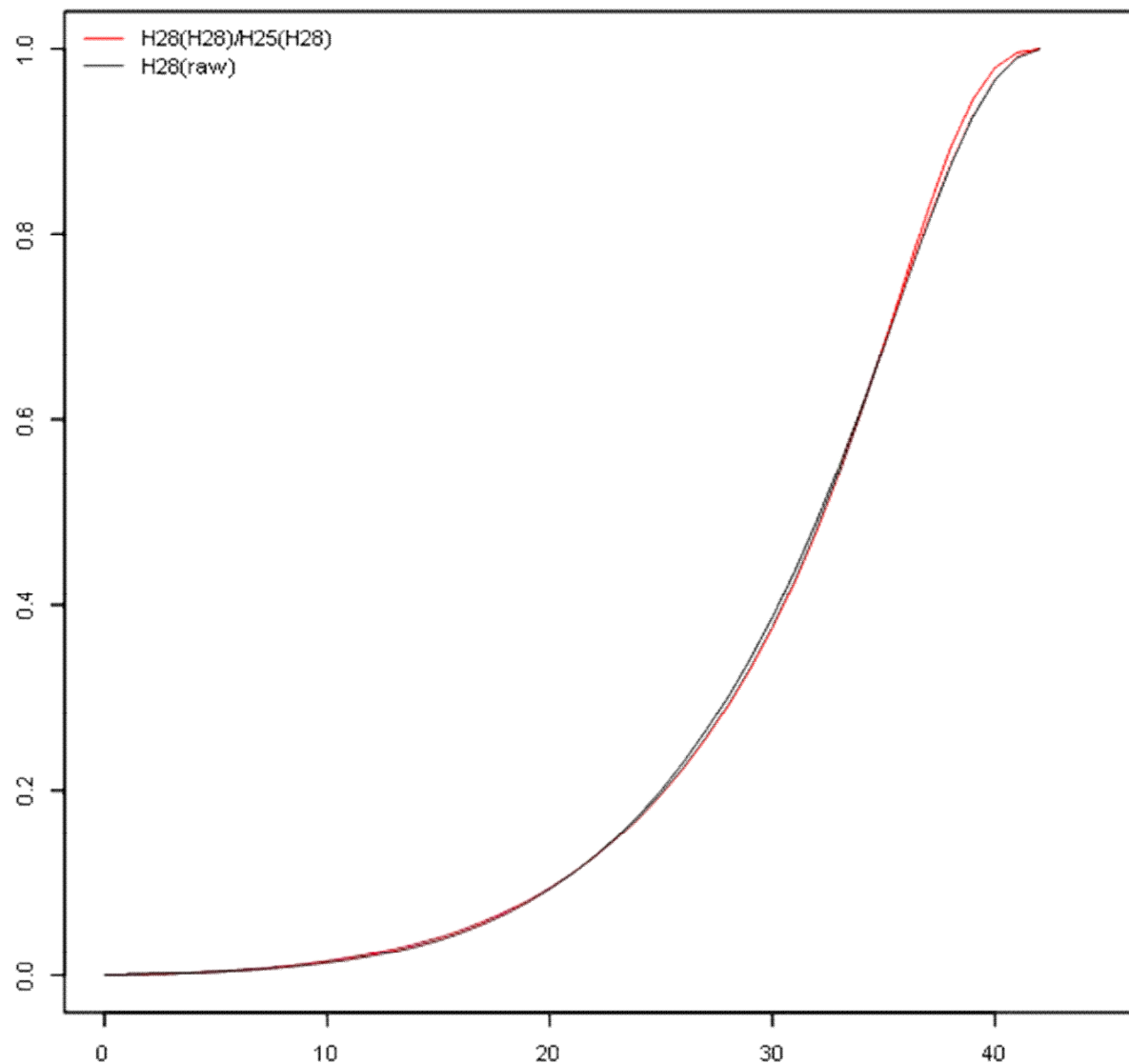
H25年度集団がH28年度調査を受けたと仮定した場合 の復元得点分布とH28年度実データ分布の比較 全国の年度間比較：小学校・国語



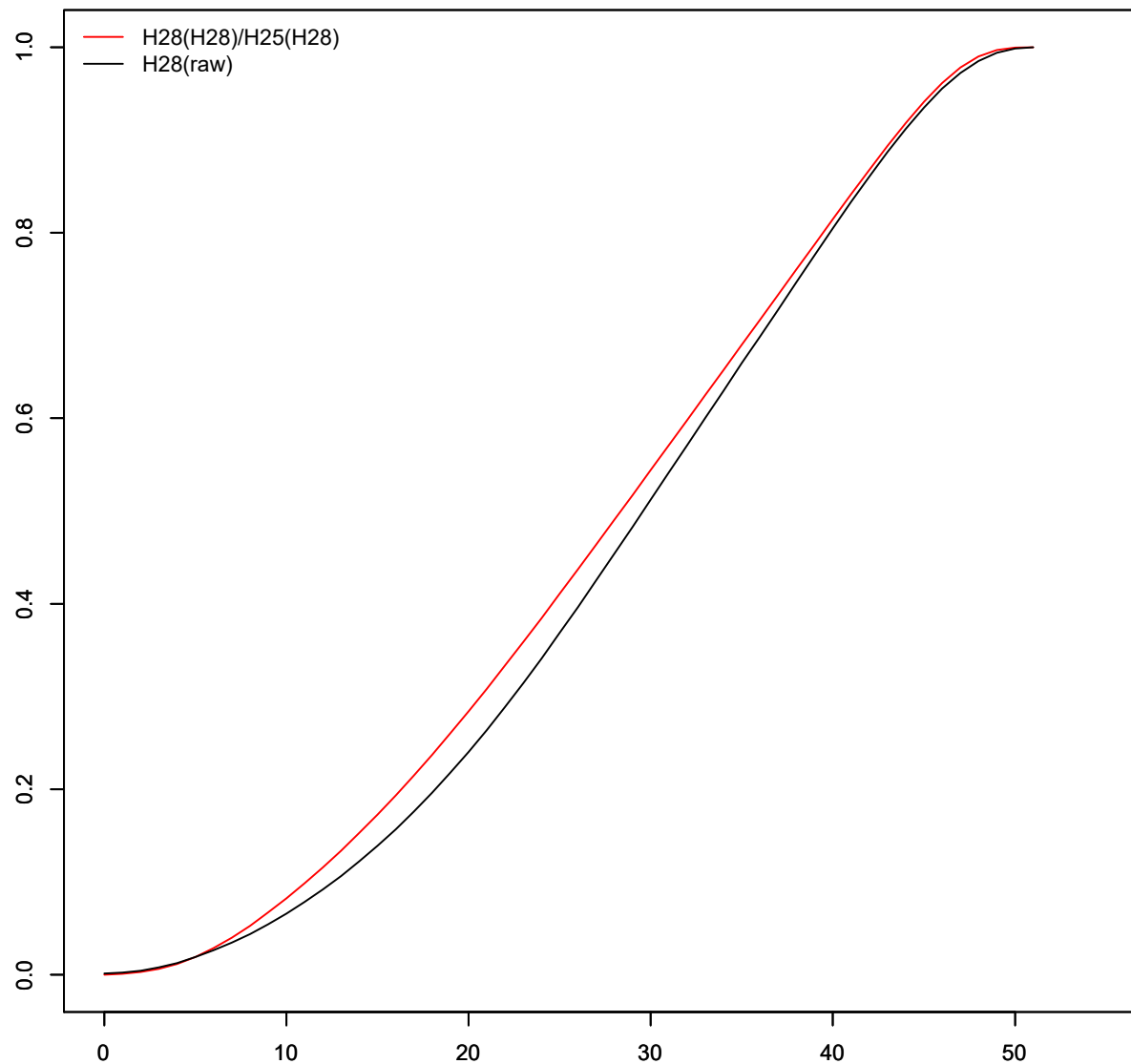
H25年度集団がH28年度調査を受けたと仮定した場合の 復元得点分布とH28年度実データ分布の比較 全国の年度間比較：小学校・算数



H25年度集団がH28年度調査を受けたと仮定した場合の
復元得点分布とH28年度実データ分布の比較
全国の年度間比較：中学校・国語

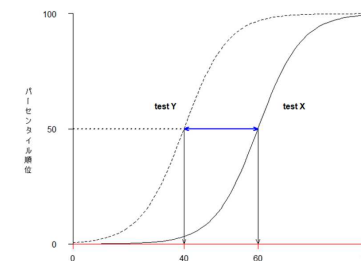


H25年度集団がH28年度調査を受けたと仮定した場合の 復元得点分布とH28年度実データ分布の比較 全国の年度間比較：中学校・数学



小学校・国語・対応表

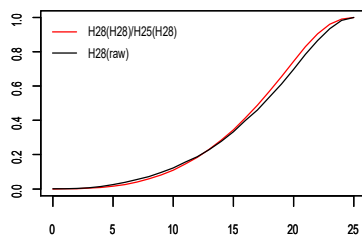
H25(raw) → H28(H28)/H25(H28)



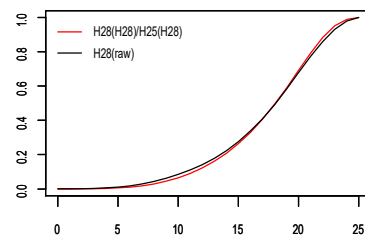
H25素得点	H28年度換算	度数
0	1.786	1,699
1	2.872	2,502
2	3.798	4,570
3	4.769	7,726
4	5.762	11,456
5	6.767	16,066
6	7.779	21,066
7	8.783	26,004
8	9.770	31,151
9	10.745	36,603
10	11.702	41,844
11	12.640	47,285
12	13.560	52,715
13	14.450	57,085
14	15.304	61,471

H25素得点	H28年度換算	度数
15	16.137	65,334
16	16.945	67,904
17	17.729	69,920
18	18.490	70,902
19	19.223	71,482
20	19.946	69,907
21	20.646	66,156
22	21.325	60,700
23	22.007	53,874
24	22.655	44,371
25	23.309	33,539
26	24.008	22,374
27	24.573	11,216
28	25.326	3,374

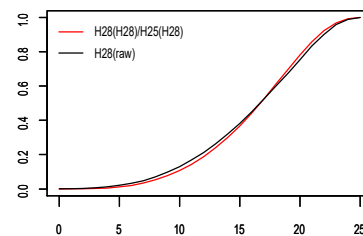
都道府県（匿名化済み）における学力分布の年度間比較の例 小学校・国語 (1/4)



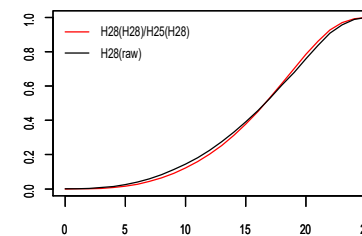
1



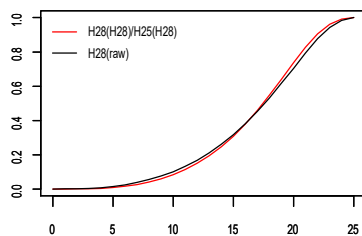
2



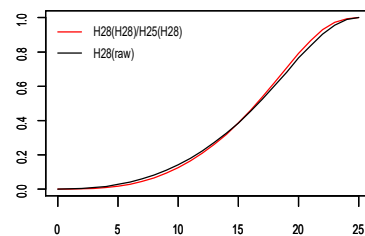
7



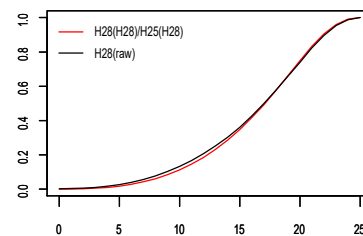
8



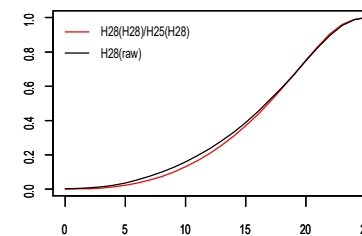
3



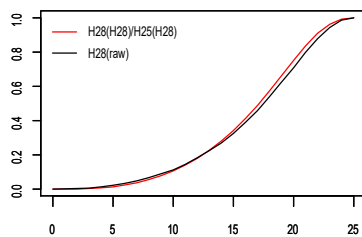
4



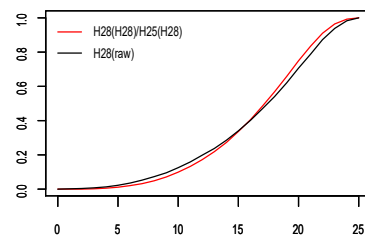
9



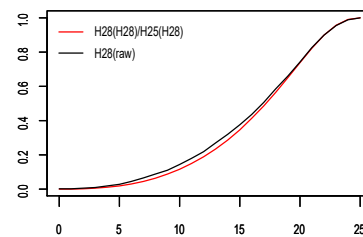
10



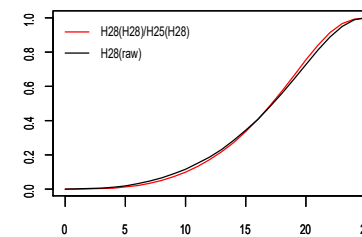
5



6

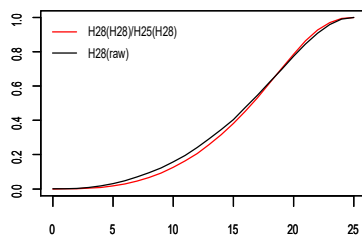


11

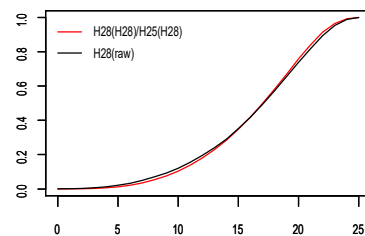


12

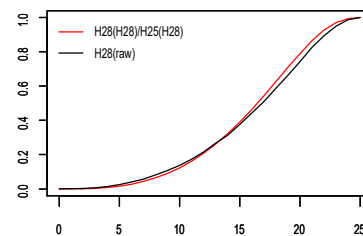
都道府県（匿名化済み）における学力分布 の年度間比較の例 小学校・国語 (2/4)



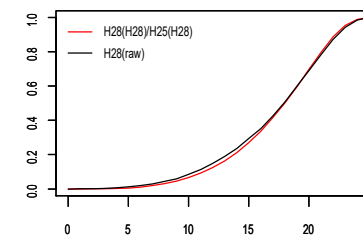
13



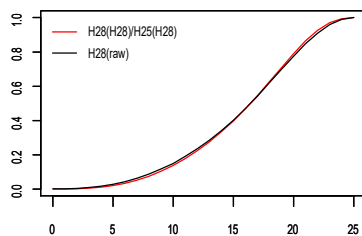
14



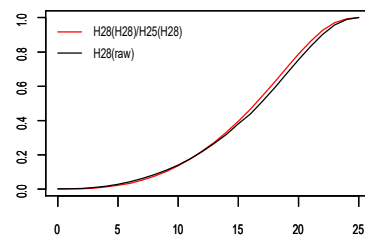
19



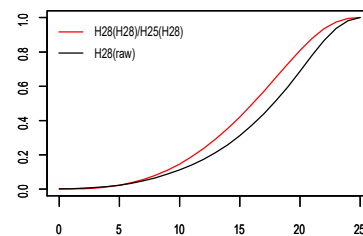
20



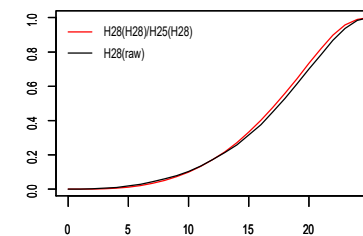
15



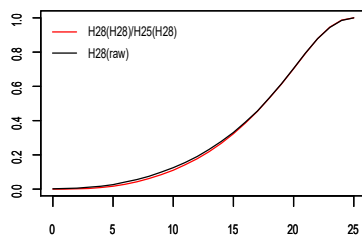
16



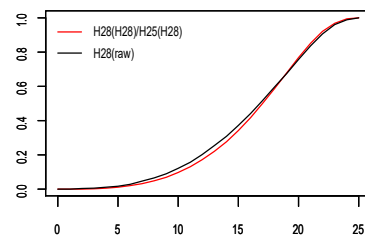
21



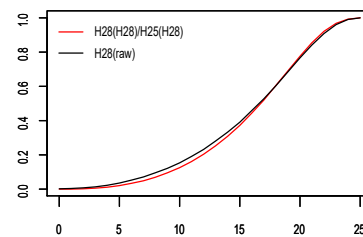
22



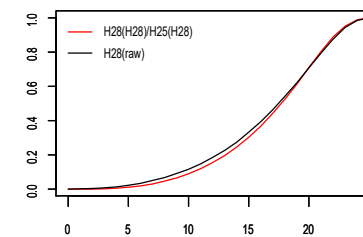
17



18

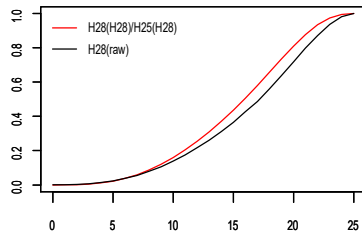


23

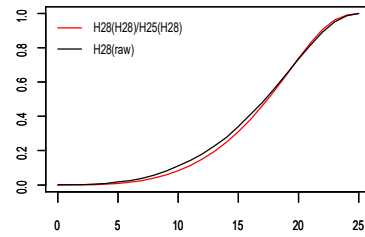


24

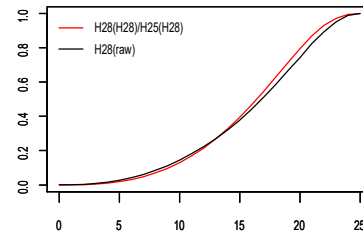
都道府県（匿名化済み）における学力分布 の年度間比較の例 小学校・国語 (3/4)



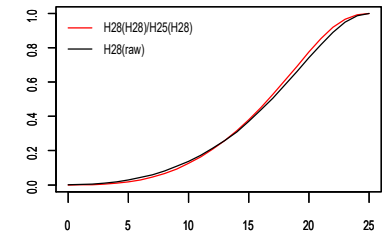
25



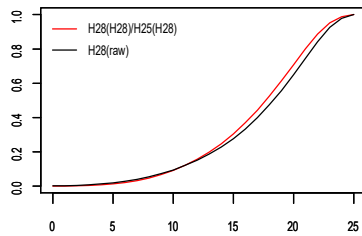
26



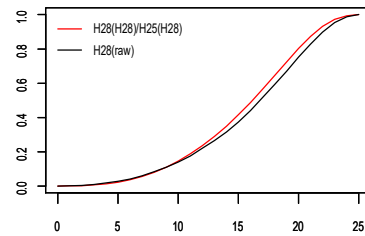
31



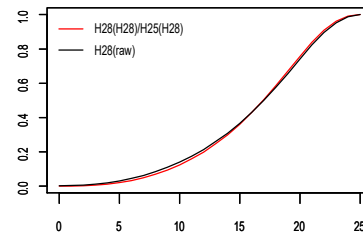
32



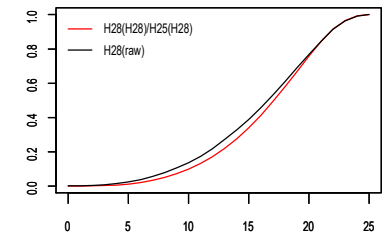
27



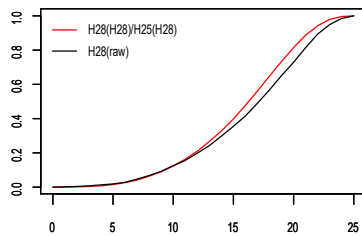
28



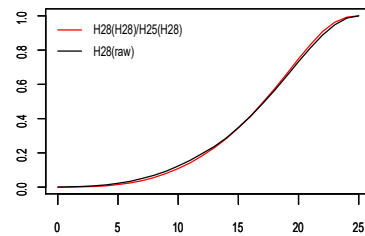
33



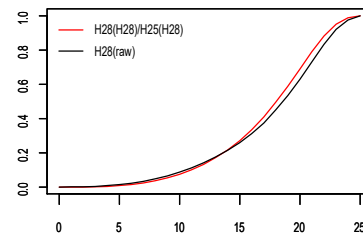
34



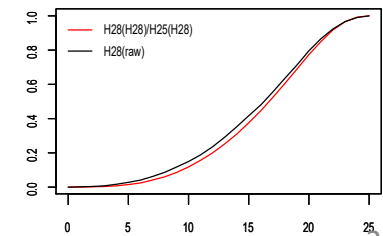
29



30

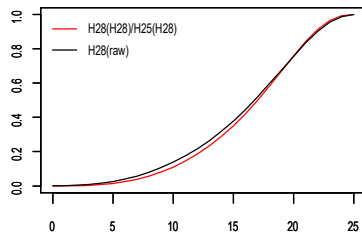


36

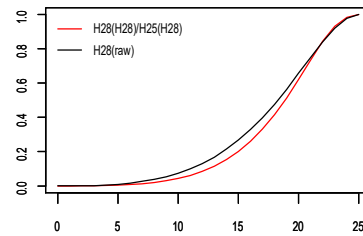


37

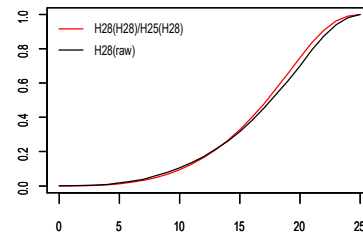
都道府県（匿名化済み）における学力分布 の年度間比較の例 小学校・国語 (4/4)



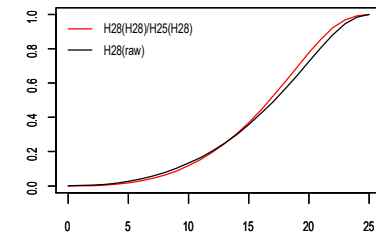
38



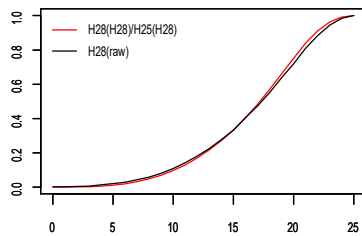
39



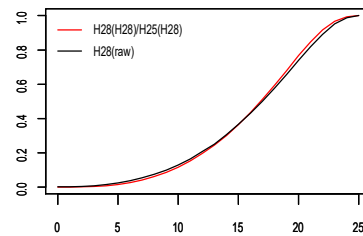
44



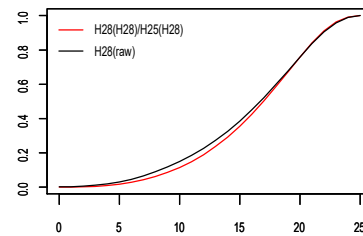
45



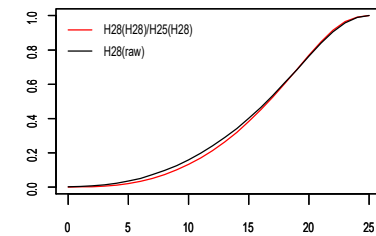
40



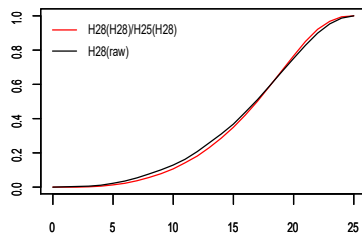
41



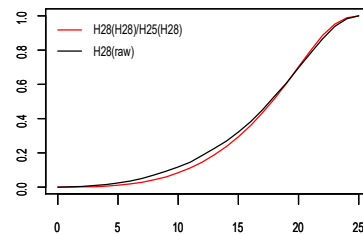
46



47



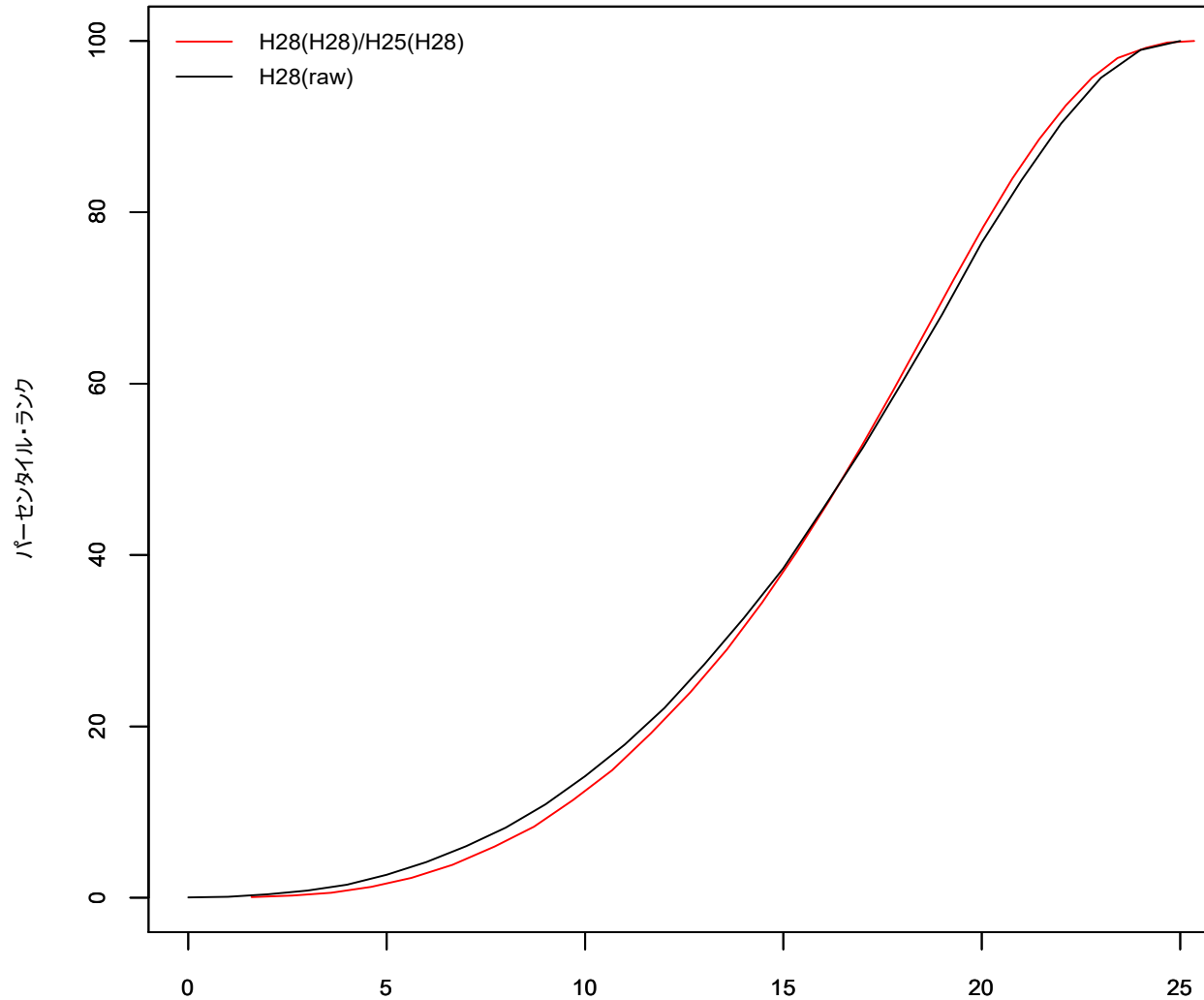
42



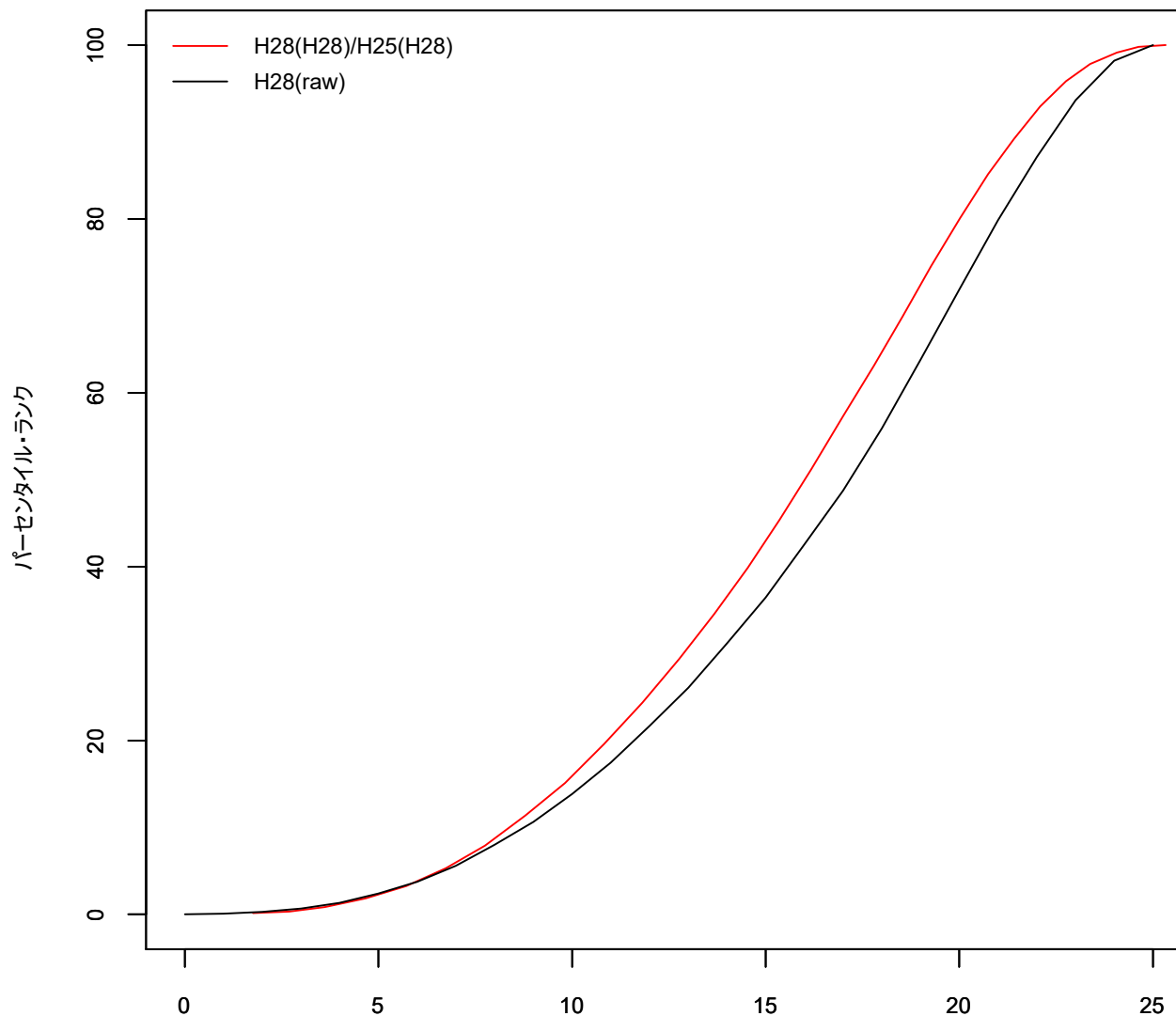
43

※1 上記は匿名化のあとランダム順に配置
※2 震災被害を受けた熊本県は含まず

匿名化第4都道府県の年度間比較 (小学校・国語)



匿名化第25都道府県の年度間比較 (小学校・国語)



3 追加提案事項：
経年変化分析調査において
推算値をもちいた下位領域ごとの
年度間比較(小学校・算数)

下位領域ごとの年度間比較の難しさ(1/2)

正答数得点(素得点)比較の限界:

- 1) 下位領域の項目数が異なれば比較できない
- 2) 項目の難しさ(困難度)が異なれば比較できない

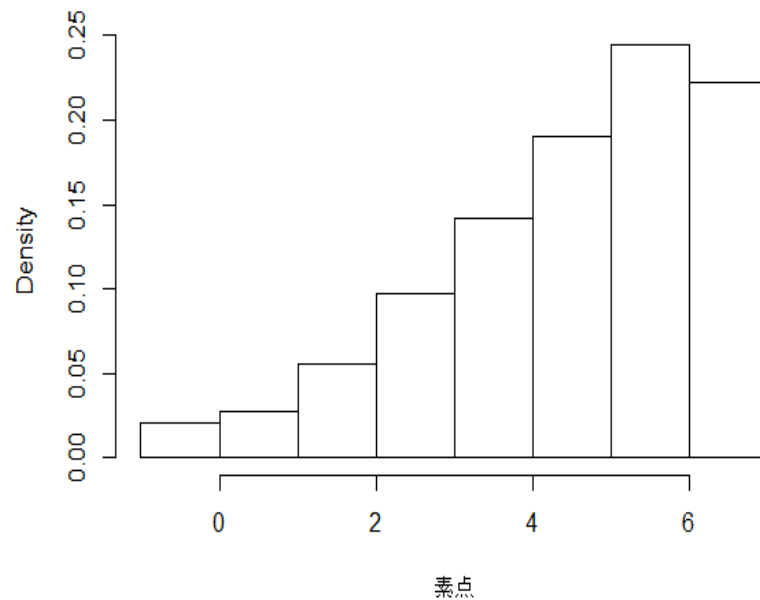


Figure 1 素点度数分布 (H25)

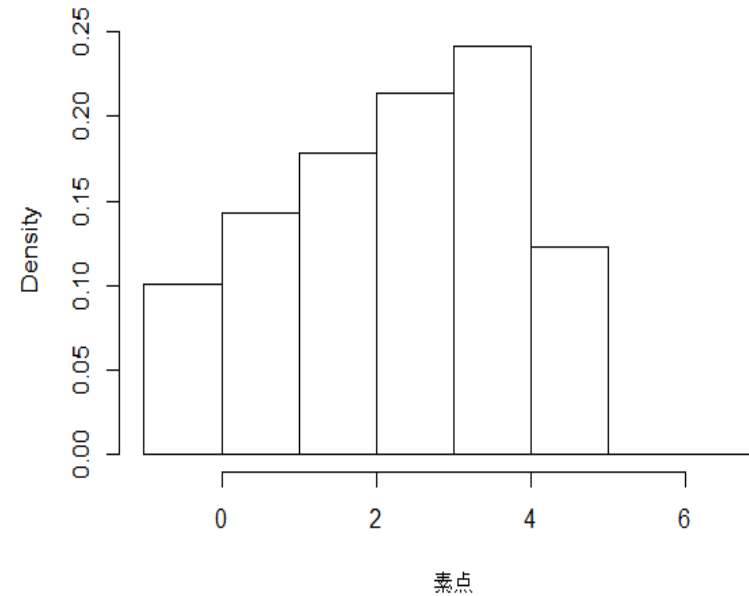


Figure 2 素点度数分布 (H28)

下位領域ごとの年度間比較の難しさ(1/2)

尺度値(IRT能力母数推定値:IRTスコア)での比較の限界:

- 1) 領域ごとのため項目数が少なく推定値が不安定
- 2) 項目数が少ないため学力の識別精度が粗くなる

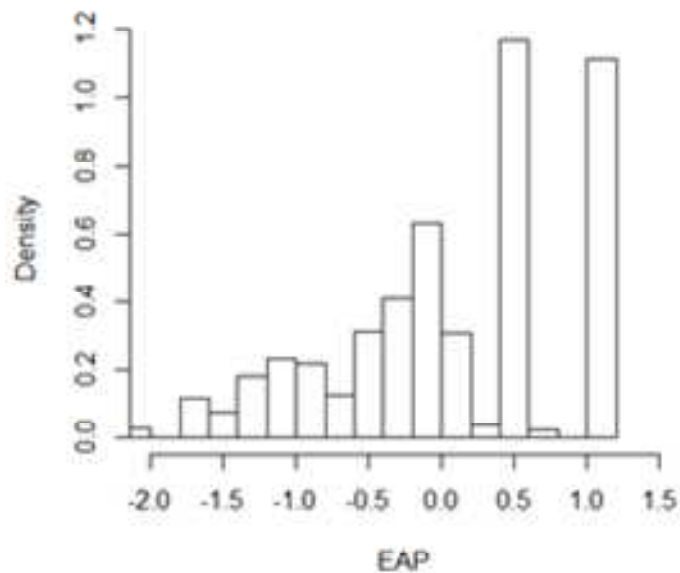


Figure 3 EAP 度数分布 (H25)

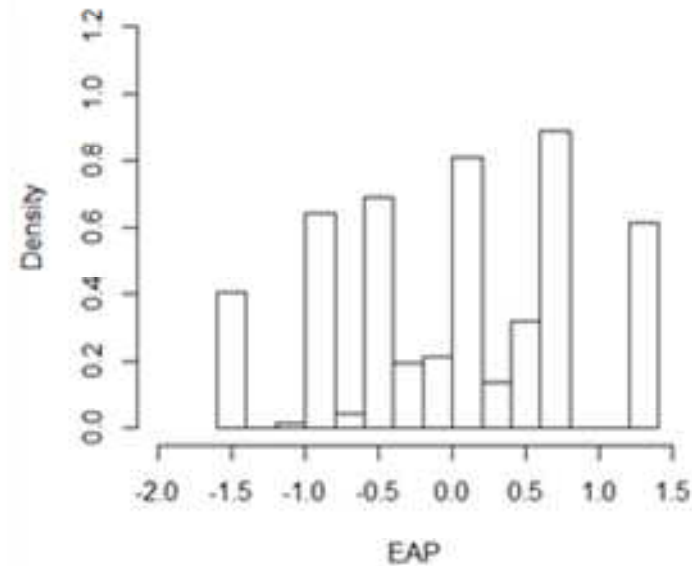


Figure 4 EAP 度数分布 (H28)

推算値利用の可能性

推算値の特徴:

- 1) 能力推定値のランダムネスを考慮できる
- 2) 少数個の項目からでも集団統計量を推定できる

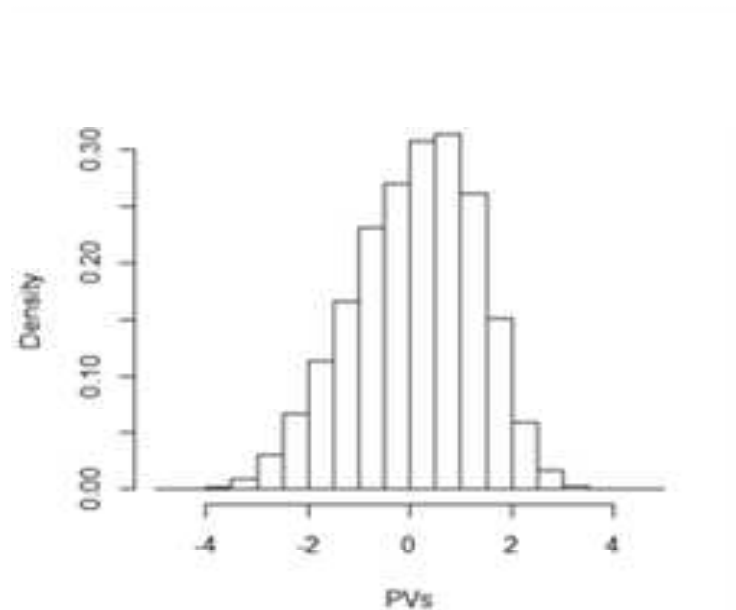


Figure 9 推算値ヒストグラム (H25)

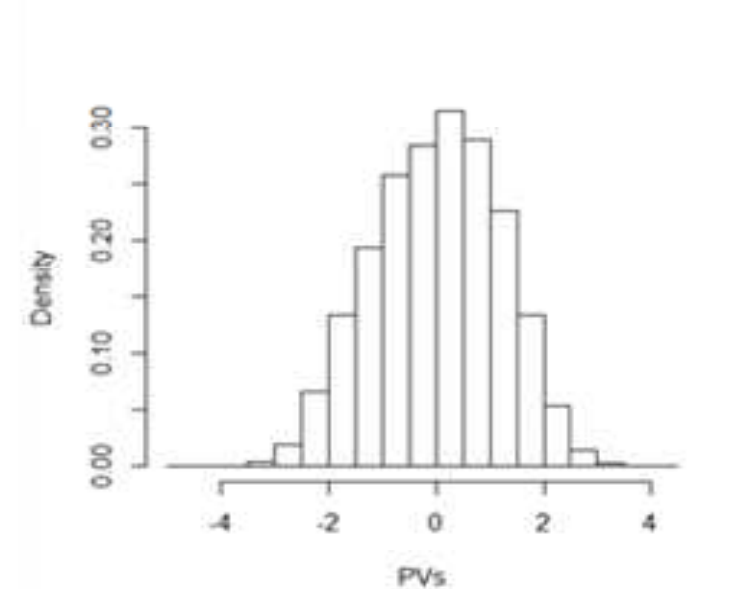


Figure10 推算値ヒストグラム (H28)

下位領域ごとの学力分布の変化

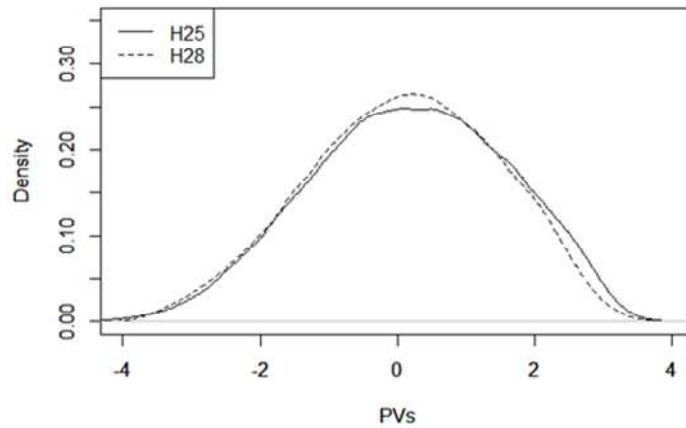


Figure 14 数と計算

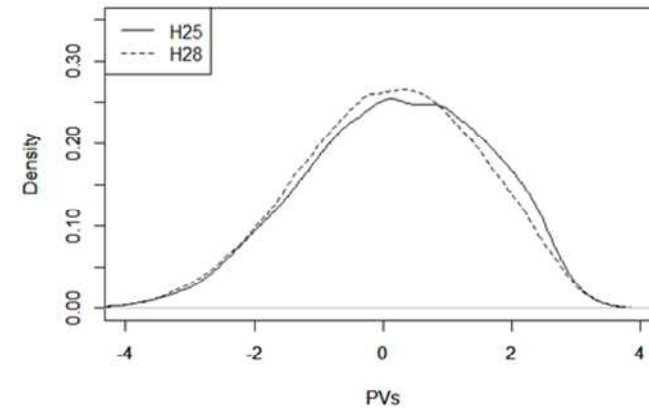


Figure 15 量と測定

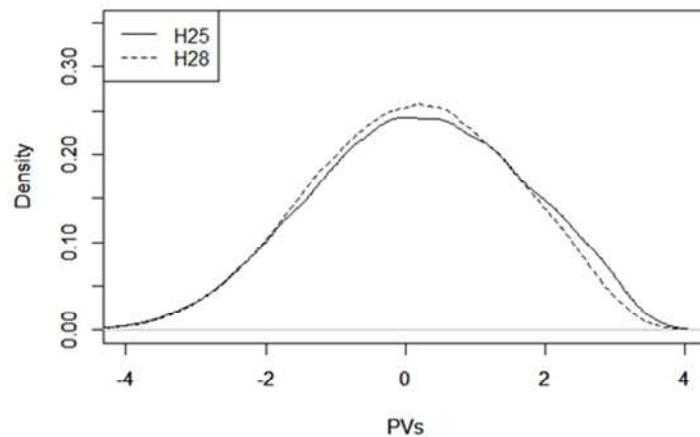


figure 16 図形

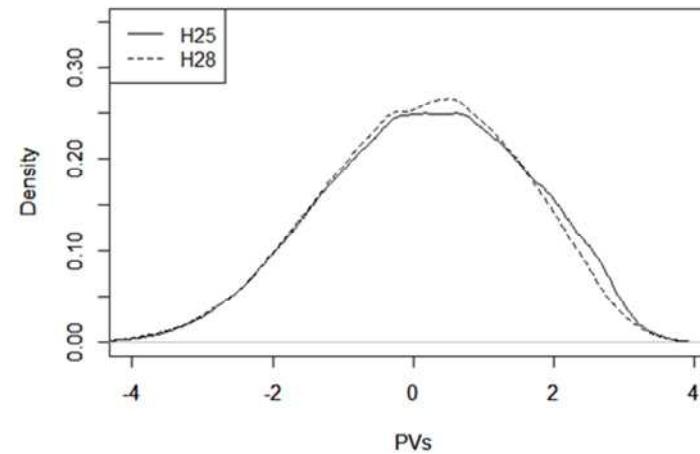


Figure 17 数量

まとめ

- 本体調査および経年変化分析調査のもつそれぞれの利点と、項目反応理論(Item Response Theory: IRT)モデルや 等化(equating)、対応づけ(linking)などの学力調査の測定技術を組み合わせることで以下の手法を開発した。
 - A) 復元得点分布(IRT observed score distribution)による年度間の学力分布の変化を捉える方法
 - B) 上記にもとづき、IRTモデルなどの複雑な専門的方法を利用しなくても、上と同等のことが各都道府県単位で実現できる、年度間の換算表(対応表: concordance table)を作成する方法
 - C) 経年変化分析調査において、項目数が少ないため安定的な経年比較ができない科目内領域ごとの年度間比較のために、推算値(Plausible Values: PVs)を利用する方法
- さらに体系的かつ詳細なシミュレーション研究により、これらの手法の精度検証が必要と考えるが、
- 全国学力・学習状況調査の目的の一つである政策立案(Evidence-Based Policy Making: EBPM)のための総合的かつ客観的根拠を供給する技術的基盤が提供できた。

