peakshift 法のロバスト性についての検討

A Study on the Robustness of the Peakshift Method

高橋 聡[*1]，北澤 正樹[*2,3]，吉川 厚[*2,4]
Satoshi TAKAHASHI[*1], Masaki KITAZAWA[*2], Atsushi YOSHIKAWA[*3]
[*1] 関東学院大学
[*1] Kanto Gakuin University
[*2] 立教大学
[*2] Rikkyo University
[*3] 北澤技研
[*3] Kitazawa Tech
[*4] 東京工業大学
[*4] Tokyo Institute of Technology
Email: satotaka@kanto-gakuin.ac.jp

あらまし：本研究では peakshift 法のロバスト性について分析を行う．peakshift 法とは，異なる試験群に対する組織ごとの合格者数に基づいて，試験難易度の順位を推定する手法である．まず，シミュレーションを利用して，評価用データを生成する．次に，組織データに欠損が生じた場合を想定し，生成したデータから段階的に組織単位でデータの削除を行う．実験結果からシミュレーションでは 20%の組織のデータが存在していれば，実データでは 30%の組織のデータが存在していれば，試験難易度の順位推定を行えることを示す．
キーワード：Peakshift，大学受験，項目反応理論

1. はじめに

本論文では，受験生が選択的に試験を受けない試験結果データに注目する．我々はこのようなデータを"selectively omitted examination data"と呼んでいる．データの例として，大学受験結果や資格試験結果，学生の就職活動結果が挙げられる．例えば大学受験結果において，受験生は自分の学力から見て入試が難しすぎたり，簡単すぎたりする大学を受験しない．さらに，受験生の学力と入試の難易度がわからない限り，受験生がどちらの理由で受験をしなかったのかを知ることができない．上記の性質により，IRT[(1)]などの既存の手法では，入試の難易度を推定することが困難である．このようなデータの試験の難易度順位を推定するために，peakshift 法が提案されている[(2,3)]．試験の難易度順位がわかれば，試験が求める能力の観点から，教育機関を評価する新たな指標を作成することができる．

本論文では，peakshift 法のロバスト性評価を目的とする．まず，シミュレーションにより仮想の受験結果データを生成する．そして，シミュレーションデータと実データの一部を段階的に削除する．最後にそれらのデータに対して peakshift 法を適用し，そのロバスト性を分析する．

2. selectively omitted examination data

図 1 は，selectively omitted examination data の例である．縦軸が各試験の合格者数，横軸が各試験の難易度を表す. また, 学力が低い組織 (例えば, 高校), 中程度の組織，高い組織ごとに合格者数がプロットされている. 各組織には, 合格者数のピークがある．

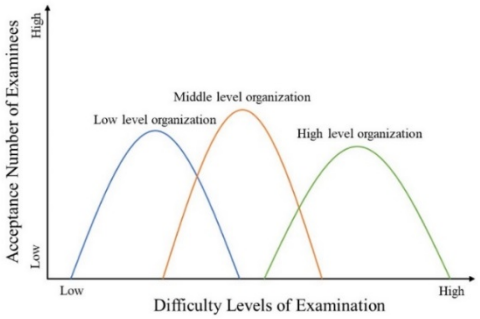

図 1 Selectively omitted examination data の例

このピークは，組織の学力が上がると，より高い難易度の試験に移動する．組織の学力を正規分布と考えると，組織の平均学力はピーク時の試験の難易度と一致する．受験する試験の難易度がピークの試験よりも高くなったり，低くなったりすると，合格者数は少なくなる．この性質により，もし，このようなデータにおいて，試験の難易度や組織の学力が事前にわからない場合，合格者数だけでは試験の難易度順位を推定することは困難となる．

IRT[(1)]などの難易度推定手法では, 誰がどの項目を解けるのか，あるいは解けないのかというデータが必要となる．しかし，selectively omitted examination data は，このデータが欠けている．受験者は，試験が簡単すぎたり難しすぎたりした場合に受験をしないが，それがどちらの理由なのかはわからないためである．

3. シミュレーションデータおよび実データ

表 1 は，調査会社から提供された首都圏の実デー

タである．各高校の生徒数，高校の数，各高校と各大学入試の合格者数, 大学の数, 各大学の合格者数，大学入試の難易度などが含まれている．生徒が合格した大学入試の難易度を，その生徒の学力を示すものとして扱った．このデータに基づいて，シミュレーションにより仮想のデータを作成した．

高校の生徒データは以下の手順で作成した．

(1) 生徒を生成し，生徒の学力を正規分布にしたがって割り当てる．
(2) 高校を生成し，高校の平均学力を正規分布にしたがって割り当てる．
(3) 生徒数を考慮して，平均学力が生徒の学力に近い高校に確率的に割り当てる．
(4) 各高校の生徒の学力の標準偏差を計算する．標準偏差が制限値を超えている高校が一つでもあれば，手順(3)に戻る．
(5) 生徒の学力の平均に基づいて，高校をランク付けする．

大学受験のデータは以下の手順で作成した．

(1) 大学を生成し，試験の難易度を正規分布にしたがって割り当てる．
(2) 各生徒は，その生徒の学力±入試能力の限界の範囲の難易度の試験をすべて受験する．
(3) 試験の難易度が高い大学順に，大学は受験生の入学判定を行う．大学は入学者数まで受験生を入学者として判定する．すでに他の大学に合格している受験生は入学させない．
(4) 各大学は，最も学力が最も低い入学者以上の学力の受験生を合格者として判定する．

4. 評価実験

シミュレーションデータと実データに対して，peakshift 法を適用し, 試験の難易度順位を推定した．そして，推定した順位と，実際の難易度の順位との間で，スピアマンの順位相関係数を算出した．

なお, peakshift 法のロバスト性を分析するために，一部の高校を削除した．削除する高校数は全高校数に対して 0%から 90%まで 10%刻みで変化させた．

表 1 シミュレーション設定

パラメータ	シミュレーション	実データ
高校生の人数	283,422	286,000
生徒の学力の平均	-	0
生徒の学力の標準偏差	-	1
高校の数	1,078	1,100
各高校の生徒数	Ave. 262.9	260
各高校の平均学力の平均	0.3	0
各高校の平均学力の標準偏差	0.64	0.8
各高校の生徒の学力の標準偏差	0.61	0.6
各高校の生徒の学力の標準偏差の限界値	-	1.96 * 0.20
大学数	157	160
各大学の入学者数	Ave. 1555.7	1,600
入学試験の難易度の平均	-	0
入学試験の難易度の標準偏差	-	1
入試能力の限界	-	1

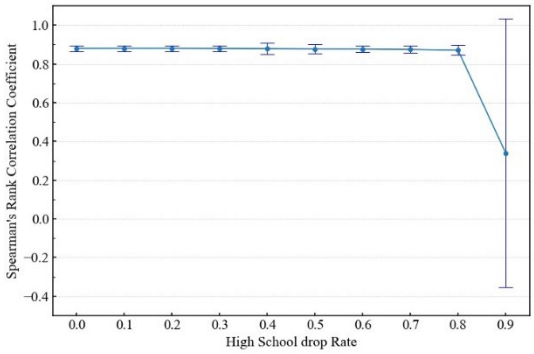

図 2 シミュレーションデータの推定結果

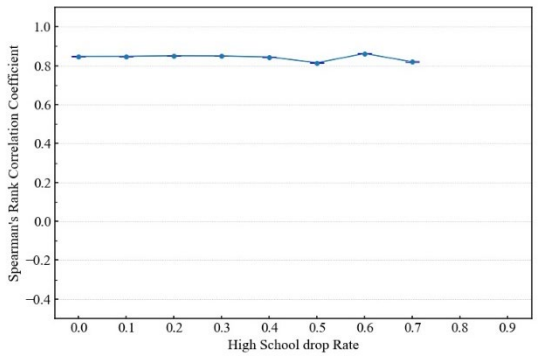

図 3 実データの推定結果

このシナリオでは，一部の組織（高校）のデータが欠落していた場合を想定している．

図 2 と 3 にシミュレーションデータと実データの推定結果を示す．図 2 から，シミュレーションでは 20%の組織のデータ(削除率 80%)が存在していれば，試験難易度の順位推定を行えることがわかる．さらに図 3 から, 実データでも 30%の組織のデータ(削除率 70%)が存在していれば, 順位推定が行えることがわかる．なお，peakshift 法では演算の中でクラスタリングを行っている．実データで削除率 80%以上の組織データにおいて, このクラスタリングが行えず，順位推定をすることができなかった．

5. おわりに

本論文では，peakshift 法のロバスト性について分析を行った．実験結果から，シミュレーションでは 20%の組織のデータが存在していれば，実データでは 30%の組織のデータが存在していれば，試験難易度の順位推定を行えることを示した．

参考文献

(1) Hambleton, R. K. and Swaminathan, H.: “Item response theory: Principles and applications”, Springer Science & Business Media, New York （2013）
(2) 青木亮磨, 北澤正樹, 高橋聡, 吉川厚, 山村雅幸: “系統的な欠損を持つデータにおける項目序列決定手法の提案” , 2019 年度 JSiSE 学生研究発表会・発表論文, pp.63-64
(3) 高橋聡, 北澤正樹, 吉川厚: “シミュレーションを利用した peakshift 法の適用範囲検討” , 2020 年度 JSiSE 全国大会