

令和 3 年 6 月 15 日現在

機関番号：82616

研究種目：若手研究

研究期間：2018～2020

課題番号：18K18166

研究課題名（和文）項目露出率の最小化のための整数計画問題を用いた等質適応型テストの提案

研究課題名（英文）Computerized Adaptive Testing Method using Integer Programming to Minimize Item Exposure

研究代表者

宮澤 芳光（Yoshimitsu, Miyazawa）

独立行政法人大学入試センター・研究開発部・助教

研究者番号：70726166

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：本研究では、整数計画問題を用いて露出数を最小化し、能力が同等な受検者であっても、同一の測定精度を保ちつつ、異なる項目を適応的に出題できる等質適応型テストを提案した。具体的には、1)露出数を最小化し、2)テストの長さとして3)テスト間の測定精度が等質であるにも関わらず、4)異なる項目を出題できる適応型テストを提案した。研究成果は、2019年にJSAI 2019 Excellence Awardを受賞し、2020年にAdvances in Artificial Intelligenceに掲載されている。

研究成果の学術的意義や社会的意義

従来の適応型テストでは、同一の受検者が複数回受験した場合には、同一の項目群が出題される傾向があり、実際に適応型テストを導入しているテスト事業者の重要な問題になっている。また、特定の項目群が過度に受検者に暴露されてしまう傾向があるため、項目内容の暴露につながり、テストの信頼性の低下要因となりうる。さらに、測定精度の等質化が実現されておらず、テストの長さや測定精度について受検者間で大きな差が生じるといふ問題がある。本研究では、これらの課題を解消し、適応型テストの実用化に寄与しており、学術的にも社会的にも意義がある。

研究成果の概要（英文）：We propose Computerized adaptive testing(CAT) item selection using integer programming. It can select a different item while maintaining equal accuracy, even for an examinee with the same ability. The approach of the proposed method is described as presented below. 1) A group that satisfies test constraints as Fisher information is assembled using integer programming for minimizing item exposure. 2) An item that has the highest information is selected from the group. The proposed method uses integer programming for minimizing item exposure and satisfying upper and lower bounds of information. Accordingly, a different item set can be selected, even for an examinee with equal ability. Moreover, the improved diversity of item selection is expected to encourage thorough use of items in an item pool and to mitigate deviation in exposure.

研究分野：教育工学

キーワード：適応型テスト eテストング 項目反応理論 教育工学

### 1. 研究開始当初の背景

近年、Web上でテストを受験する CBT (Computer based testing) を用いた e テスティングが実用化されつつある。e テスティングでは、受検者が何度でも等質な測定精度で異なる項目から構成されたテストを受検できる。このため、テストの結果が受検者に大きな影響を与えるハイ・ステークスなテストで導入が検討されている。e テスティングや CBT の技術として適応型テスト (CAT: Computerized Adaptive Testing) と呼ばれるテスト出題方式が知られている。適応型テストでは、受検者が項目を解答するたびに能力を推定し、その能力推定値に対して情報量の高い項目を出題する。これにより、測定精度を減少させずに、出題項目数や受検時間を減らすことができる。しかし、同一の受検者が複数回受験した場合には、同一の項目群が出題される傾向があり、実際に適応型テストを導入しているテスト事業者の重要な問題になっている。さらに、特定の項目群が過度に受検者に露出されてしまう傾向があるため、項目内容の暴露につながり、テストの信頼性の低下要因となりうる。適応型テストを運用するためには、難易度や識別力などの特性が既知のテスト項目集合 (アイテムバンクと呼ばれる) を事前に集積する必要があり、特にハイ・ステークスなテストにおいては、項目の作成に膨大な経済的・時間的コストを要するため、テストの実施においてはアイテムバンク内の全ての項目をできる限り活用することが望ましい。これまでに、特定の項目の過度な露出を防ぐため、項目の選択に対して制約を付けた適応型テストが提案されている。しかし、これらの研究では、測定精度の等質化が実現されておらず、テストの長さや測定精度について受検者間で大きな差が生じるという問題が指摘されていた。

### 2. 研究の目的

本研究では、従来の適応型テストの課題を解決するため、整数計画問題を用いて項目の露出を最小化し、能力が同等な受検者であっても、同一の測定精度を保ちつつ、異なる項目を適応的に出題できる等質適応型テストを提案する。具体的には、1)露出数を最小化し、2)テストの長さとして3)テスト間の測定精度が等質であるにも関わらず、4)異なる項目を出題できる適応型テストを提案する。

提案手法の具体的なアプローチは以下の通りである。

1. 受検者への項目選択のたびに、整数計画問題を用いて測定精度の上限・下限を満たしながら露出数が最小になる項目集合を構成する。
2. 項目集合から測定精度が最も高い項目を選ぶ。

提案手法では、アイテムバンク内の項目を満遍なく利用でき、露出数を減少させることができることが期待され、能力が同等な受検者であっても異なる項目群を出題することができる。特に、資格検定などのハイステークスなテストでは、アイテムバンクを有効活用することができるため、項目の露出数を減少させ、効果的な項目の活用が期待される。

### 3. 研究の方法

本研究では、露出数を最小化する等質適応型テストを定式化し、シミュレーション実験・実データを用いた実験から本手法の有効性を明らかにする。まず、露出数を最小化した等質適応型テストにおける項目選択を定式化する。次に、項目反応理論に基づいて項目の特性 (難しさ等) を乱数から作成し、アイテムバンクを構築した後、本手法を用いて適応型テストのシミュレーションを実施する。シミュレーション実験では、本手法の限界や適用条件について評価するためにいくつかの実験条件を設定し、提案手法の有効性を検討する。さらに、実データを用いて提案手法の有効性を評価する。具体的には、実際のテスト業者のアイテムバンクを用いて、シミュレーション実験と同様に本手法の有効性を検証する。

### 4. 研究成果

ここでは、本研究で得られた成果として「露出数を最小化するための整数計画問題を用いた等質適応型テスト」の手法について紹介する。本手法のアルゴリズムは、以下の通りである。

1. 受検者の能力値を初期化する。
2. 整数計画問題を用いて条件を満たす項目群を構成する。ここでは、整数計画問題を用いたテスト構成手法の一つである Adema [1] らの手法を拡張し、目的関数に項目の露出数の最小化を組み込む。 $k = 1, 2, \dots, K$  での  $\theta_k$  における情報量の下限は、 $r_k$  とする。また、 $\theta_k$  における情報量の上限は、 $s_k$  として表す。 $\theta_k$  における情報量は、 $I(\theta_k)$  と示す。また、項目  $i$  の露出数を  $e_i$  とする。アイテムバンクの項目  $i = 1, \dots, I$  における  $x_i$  は、1 なら項目群に含まれることを表し、0 ならば項目群に含まれないことを示す。以下の整数計画問題を用いて能力測定精度の上限・下限を満たしながら露出数の合計が最小になる項目集合を構成する。

$$\text{Minimize } y = \sum_{i=1}^I e_i x_i$$

Subject to

$$\sum_{i=1}^I I(\theta_k)x_i \geq r_k$$

$$\sum_{i=1}^I I(\theta_k)x_i \leq s_k$$

$$\sum_{i=1}^I x_i = n$$

3. 上記の項目集合から測定精度が最も高い項目を選ぶ。
4. 受検者は項目に解答する。
5. 受検者の解答を自動採点し、その結果と解答履歴から受検者の能力を測定する。
6. 終了条件を満たしていれば、終了、そうでなければ項目の出題を繰り返す。

シミュレーションから生成したアイテムバンクと実際のテスト事業者で運用されているアイテムバンクを用いて提案手法の有効性を検証した。これらの結果から次のことがわかった。(1) 項目の露出数を小さくすることができた(2)テストの長さの等質性を高めることができた(3) 受検者に出題したテスト間で同一ではない項目の割合を表す「異なる項目の割合」が高い値を示した。

上記の提案手法は、図1の通り、e テスティングシステムに実装されており、実際にテストを受験することができる。

#### 参考文献

[1] Adema, J. J.: Implementations of the branch-and-bound method for test construction problems. Methodika 6(2), 99-117 (1992)

eTesting System

1項目 / 全6項目

以下の証明はどこでおかしくなったか?

$$\frac{1}{8} > \frac{1}{4}$$

証明

$$3 > 2$$

両辺に $\log_{10}\left(\frac{1}{2}\right)$ をかける

$$3 \log_{10}\left(\frac{1}{2}\right) >_{(1)} 2 \log_{10}\left(\frac{1}{2}\right)$$

両辺の $\log_{10}\left(\frac{1}{2}\right)$ の係数を真数の肩にかける

$$\log_{10}\left(\frac{1}{2}\right)^3 >_{(2)} \log_{10}\left(\frac{1}{2}\right)^2$$

両辺を真数に戻す

$$\left(\frac{1}{2}\right)^3 >_{(3)} \left(\frac{1}{2}\right)^2$$

$$\frac{1}{8} >_{(4)} \frac{1}{4}$$

選択肢(1)の表示・非表示

選択肢			
<input type="radio"/> $3 \log_{10}\left(\frac{1}{2}\right)$	<input type="radio"/> $2 \log_{10}\left(\frac{1}{2}\right)$	<input type="radio"/> $\log_{10}\left(\frac{1}{2}\right)^3$	<input type="radio"/> $\log_{10}\left(\frac{1}{2}\right)^2$
<input type="radio"/> $\left(\frac{1}{2}\right)^3$	<input type="radio"/> $\left(\frac{1}{2}\right)^2$	<input type="radio"/> $\frac{1}{8}$	<input type="radio"/> $\frac{1}{4}$
<input type="radio"/> $>_{(1)}$	<input type="radio"/> $>_{(2)}$	<input type="radio"/> $>_{(3)}$	<input type="radio"/> $>_{(4)}$

解答

図1 e テスティングシステムを用いたテストの画面例

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Yoshimitsu MIYAZAWA, Maomi Ueno	4. 巻 1128
2. 論文標題 Computerized Adaptive Testing Method Using Integer Programming to Minimize Item Exposure	5. 発行年 2020年
3. 雑誌名 Advances in Artificial Intelligence	6. 最初と最後の頁 105-113
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-030-39878-1_10	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 虻名 哲也, 宮澤 芳光, 森本 康彦	4. 巻 37
2. 論文標題 項目反応理論に基づく学習状況可視化システムの開発と評価	5. 発行年 2020年
3. 雑誌名 教育システム情報学会誌	6. 最初と最後の頁 317-329
掲載論文のDOI（デジタルオブジェクト識別子） 10.14926/jsise.37.317	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 宮澤 芳光, 宇都 雅輝, 石井 隆稔, 植野 真臣	4. 巻 J101-D
2. 論文標題 測定精測定精度の偏り軽減のための等質適応型テストの提案	5. 発行年 2018年
3. 雑誌名 電子情報通信学会論文誌	6. 最初と最後の頁 909-920
掲載論文のDOI（デジタルオブジェクト識別子） 10.14923/transinfj.2017LEP0028	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計7件（うち招待講演 1件 / うち国際学会 1件）

1. 発表者名 宮澤芳光, 植野真臣
2. 発表標題 複数等質テスト構成技法を用いた適応型テストの項目選択手法の提案
3. 学会等名 日本行動計量学会
4. 発表年 2020年

1. 発表者名 宮澤芳光, 植野真臣
2. 発表標題 Uniform adaptive testing using maximum clique algorithm
3. 学会等名 第19回情報科学技術フォーラム FIT2020, 情報処理学会 (招待講演)
4. 発表年 2020年

1. 発表者名 Maomi Ueno, Yoshimitsu Miyazawa
2. 発表標題 Uniform adaptive testing using maximum clique algorithm
3. 学会等名 Artificial Intelligence in Education (国際学会)
4. 発表年 2019年

1. 発表者名 宮澤芳光, 植野真臣
2. 発表標題 最適化問題を用いた等質適応型テスト
3. 学会等名 日本行動計量学会
4. 発表年 2019年

1. 発表者名 Yoshimitsu Miyazawa, Maomi Ueno
2. 発表標題 Computerized Adaptive Testing Method using Integer Programming to Minimize Item Exposure
3. 学会等名 人工知能学会
4. 発表年 2019年

1. 発表者名 宮澤芳光, 植野真臣
2. 発表標題 項目暴露率を考慮した等質適応型テスト
3. 学会等名 日本テスト学会
4. 発表年 2019年

1. 発表者名 宮澤 芳光, 植野 真臣
2. 発表標題 整数計画問題を用いた等質適応型テストの提案
3. 学会等名 日本テスト学会
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関