

テスト理論に基づいた作問アドバイス生成システムの開発

Development of an Advice Generation System for Creating Test Questions Based on the Test Theory

林 貴史^{*1}, 高木 正則^{*2}, 山田 敬三^{*2}, 佐々木 淳^{*2}

Takafumi HAYASHI^{*1}, Masanori TAKAGI^{*2}, Keizo YAMADA^{*2}, Jun SASAKI^{*2}

^{*1} 岩手県立大学大学院ソフトウェア情報学研究科

^{*1} Graduate School of Software and Information Science, Iwate Prefectural University

^{*2} 岩手県立大学ソフトウェア情報学部

^{*2} Faculty of Software and Information Science, Iwate Prefectural University

Email: g231n028@s.iwate-pu.ac.jp

あらまし：多くの教育現場では、学習者の理解度や能力を測定するためにテストが実施されている。しかし、テスト作成者が出題した問題の良し悪しを評価することは少ない。その要因として、テスト問題の分析に必要なテスト理論や統計などの専門知識が不足していることや、分析に必要な時間の確保が困難なことが考えられる。本研究では、テスト理論などの知識を持たないテスト作成者の作問支援を目的とし、テスト問題の分析結果や次回作問時のアドバイスを提示するシステムを開発した。

キーワード：テスト理論, 検定試験, 多肢選択式問題, ご当地検定, 作問支援システム

1. はじめに

多くの教育現場では、テストの得点から各学習者を評価するが、テストに出題された問題を評価することは少ない。その要因として、テスト問題の分析に必要なテスト理論や統計などの専門知識が不足していることや、分析に必要な時間の確保が困難なことが考えられる。また、テスト問題の分析結果から、作問時の改善点を導き出すのも難しい。そこで、我々はテスト問題の分析負担の軽減を目的とし、テスト問題の分析支援システムを提案する。本システムでは、テスト問題の分析結果や次回作問時のアドバイスを生成することによって、テスト理論の専門知識を持たないテスト作成者を支援する。

本稿では、岩手県盛岡市で開催されているご当地検定「盛岡もの識り検定⁽¹⁾」(以下、もりけん)の被験者データを用いて、分析結果の提示方法や作問アドバイスの生成ルールを検討する。

2. 関連研究

樋口は⁽²⁾、テスト理論の知見を有さない教授者が容易に使用できるテスト結果分析 Web アプリケーションを開発している。また、熊谷は⁽³⁾、項目反応理論によるテスト分析を行う無料ソフトウェアの Easy Estimation を開発している。これらは、テスト結果の分析といった点で本研究と類似するが、次回作問時の作問アドバイスの生成までは対象としていない。

3. 研究課題

本システムでは、テスト理論の知識を持たないテスト作成者でも理解できるテスト分析結果の表示方法や、次回作問時に参考になる作問アドバイスを生成するためのルールを作成することを狙いとしている。作問アドバイス生成ルールは古典的テスト理論や項目反応理論によって算出される項目困難度や項

目弁別力、さらには信頼性を表すクロンバックの α 係数などの値に基づいて生成する。

4. 手動によるテスト問題の分析

作問アドバイス生成ルールを構築するために手動でテスト問題を分析し、分析結果から作問アドバイスの生成手法を検討した。

4.1 分析手法

一般的にテストは信頼性と妥当性によって評価される。信頼性はクロンバックの α 係数、折半法、再テスト法、平行テスト法、など数多くの手法によって評価される⁽⁴⁾。今回の分析では、一般的に広く用いられ、テスト全体の信頼性係数を求められるクロンバックの α 係数を用いた。妥当性に関しては、数理的に求める事が困難な為、今回は評価指標としては除外した。また、項目難易度と項目弁別力を古典的項目分析手法で算出した。表 1 にテスト問題の評価指標と評価手法、評価基準を示す。

表 1：テスト問題の評価指標

指標	手法	基準
信頼性	再テスト法	0.80 以上であれば、信頼性が高い ⁽⁵⁾
	平行テスト法	
	折半法	
	内部一貫法 (クロンバックの α)	0.72 以上であれば、信頼性が高い ⁽⁶⁾
難易度	古典的テスト理論 項目反応理論	0.39 以下または 0.81 以上の項目は難易度が不適切、0.29 以下または 0.91 以上の項目は難易度が極端に高い・低い ⁽⁴⁾
弁別力	古典的テスト理論 項目反応理論	0.29 以下の項目は弁別力が低く不十分、0.19 以下の項目は弁別力が極端に低い ⁽⁴⁾

4.2 分析対象データ

2012年に実施されたもりけん3級(100問)の全受検者120名分の解答データを活用して分析した。解答データは、もりけんを主催する盛岡商工会議所(秘密保持契約を締結して共同研究を実施中)から入手した。

4.3 分析結果

上記分析対象データに対してクロンバックの α 係数を求めた結果、0.911となり、表1の基準からみて信頼性が高いテストであったことが分かる。古典的項目分析の項目難易度、項目弁別力で分析した結果の一部を表2に示す。表2の難易度、弁別力で下線が引かれている項目は表1の評価基準で「不適切・不十分」と判定された項目である。また、二重下線が引かれている項目は「極端に値が高い・低い」と判定された項目である。そのため、下線や二重下線が引かれている項目に関しては、項目の削除および修正が必要であると言える。

表2：項目分析結果

項目番号(仮)	難易度	弁別力
1	<u>0.92</u>	<u>0.19</u>
2	<u>0.95</u>	<u>0.03</u>
3	<u>0.94</u>	<u>0.09</u>
4	<u>0.96</u>	<u>0.16</u>
.....
97	0.58	0.34
98	0.57	0.38
99	0.80	0.44
100	<u>0.81</u>	<u>0.25</u>

4.4 作問アドバイス生成モジュール

以上の分析結果に基づき、作問アドバイスの生成ルールを作成した。表3に作問アドバイス生成ルールの一例を示す。今後さらなる分析を繰り返して、より適切な作問アドバイスルールを作成する。

表3：作問アドバイス生成ルール例

条件	作問アドバイス例
難易度 0.91以上 かつ 弁別力 0.19以下	・かなり易しい問題です ・成績に全く関係なく解ける問題で、合否を判定するのにあまり役立っていない問題です ・問題文の問い方を難しくするか、正答とより関連の深い誤答選択肢を設定し、難易度が上がるように改善する必要があります
難易度 0.29以下 かつ 弁別力 0.19以下	・かなり難しい問題です ・受験者は偶然に正解した可能性が高いです(多肢選択形式問題の場合) ・成績に全く関係なく解ける問題で、合否を判定するのにあまり役立っていない問題です ・問題文の問い方を易しくするか、正答とより関連の低い誤答選択肢を設定し、難易度が下がるように改善する必要があります

5. 作問アドバイス生成システム

5.1 システム概要

システムの概要図を図1に示す。テスト作成者は被験者の解答データを本システムに入力し(図1①)、項目分析モジュールで古典的テスト理論や項目反応理論などを駆使し、各テスト問題を分析する。分析

結果は利用者にフィードバック(図1②)されるとともに、作問アドバイス生成モジュールに入力され、アドバイス生成ルールDBを参照して次回作問する際に参考になる作問アドバイスを提示する(図1③)。

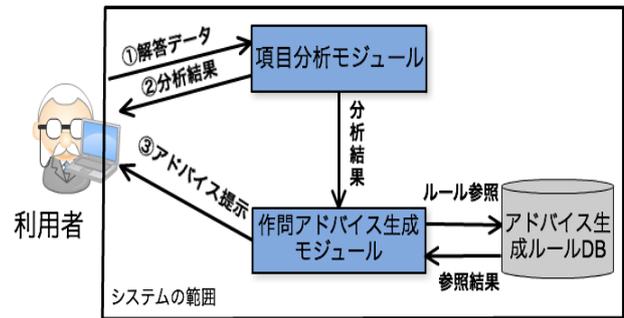


図1：システム概要図

5.2 作問アドバイス生成システムの開発

利用者自身が作成した各問題に対する作問上のアドバイスを提示し、テスト問題の分析支援と次回以降の作問を支援するシステムを開発した。図2はシステム利用者がテスト結果を入力し、分析結果と作問アドバイスを受け取る画面例である。図2左は古典的テスト理論を使用した分析結果を示し、図2右は各テスト問題の作問アドバイスである。



図2：テスト分析結果と作問アドバイス生成画面例

6. まとめと今後の課題

本稿では、テスト問題の分析結果から作問アドバイスを生成するシステムを開発した。作問アドバイスの生成ルールは、もりけんのテスト結果を手動で分析して検討した。今後は、項目難易度と項目弁別力以外から作問アドバイスを生成するルールについても検討し、作問現場での導入を目指す。

参考文献

- (1) 盛岡商工会議所, 盛岡もの識り検定 <http://www.ccimorioka.or.jp/jinzai/moriken.htm>
- (2) 樋口三郎:『テストおよびアイテム分析 Web サービスの開発』, 教育システム情報学会第39回全国大会講演論文集, pp.377-378, 2014.
- (3) 熊谷龍一: 項目反応理論と EasyEstimation., <http://irtanalysis.main.jp>, 2011.
- (4) 大友賢二:『言語テスト・データの新しい分析法 項目応答理論入門』, 大修館書店, 1996.
- (5) 山森光陽, 前田啓朗(編)『英語教師のための教育データ分析入門』, pp.4-12, 東京:大修館, 2004.
- (6) Nunnally, JumC. Psychometric Theory 2nd Edition. New York: McGraw-Hill, 1978.