

テストおよびアイテム分析 Web サービスの開発

Development of a Test and Item Analysis Web Service

樋口 三郎^{*1}Saburo HIGUCHI^{*1}^{*1} 龍谷大学理工学部^{*1} Faculty of Science and Technology, Ryukoku University

Email: hig@math.ryukoku.ac.jp

あらまし：科目内におけるテストは、各学習者の状況、クラスの状況、問（アイテム）の性質など有用な情報を含む。しかし、成績が表計算ソフトウェアに記録されるような運用の科目では、平均点、分散、ヒストグラム程度の情報しか利用されないことが多い。単純なフォーマットのデータを入力としてテストを分析しレポートを出力する、広い範囲の科目で利用可能な Web システムの開発を報告する。このシステムは、古典テスト理論等で分析を行い、様々な用途のレポートを生成する。

キーワード：テスト、アイテム分析、古典テスト理論、学習評価、Web

1. はじめに

科目内におけるテストは、学習目標への到達を点数で評価するために行われる。したがって、点数の平均値、分散、最大値、最小値などの統計量や分布には大きな関心が向けられる。しかし、テストの実施結果にはそれよりも多くの情報が含まれている。

多くの場合、テストは複数の問（項目、アイテム）からなり、テストの点数は項目の点数の合計とされる。項目 N 個の各々が 0 点以上 A 点以下の整数で評価される AN 点満点のテストを考えると、テストの合計の点数は $0, 1, \dots, N$ の $(N+1)$ 通りだが、受験者の得点のパターンは $(A+1)^N$ 通り存在する。この、得点パターンに含まれる追加の情報は、各項目の性質、各受験者の性質を含む。これらの情報を得点のパターンから抽出する方法として、古典テスト理論、項目応答理論がある⁽¹⁾⁽²⁾⁽³⁾。

得点パターンの分析を実装している LMS は多い。例えば Moodle は、1 個の小テストを構成する問題群に対して古典テスト理論に基づく分析を行い教授者に表示する。しかし、大学の科目のテストにおいて、得点パターンに着目した分析を教授者が系統的に行っていることは多くないようである。その理由は、紙のテストで得た成績を表計算ソフトウェアで管理する運用では、そのような分析を行うための簡易なツールが一般的でないためであると推測する。分析のためのツールとして、Windows 上で動作するスタンドアローンアプリケーション Exametrika、EasyEstimation や、R のパッケージ CTT、irtoys、項目応答理論に基づく商用ソフトウェア BILOG、MULTILOG などが存在するが、いずれも教授者の学習コストを要する。また、Moodle の古典テスト理論による分析にしても、LMS 外部で生成したデータに対しては適用できず、汎用性に欠ける。

本研究では、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、他システムと疎に結合された、テスト結果分析 Web アプリケーションを開発した。

2. 古典テスト理論

複数の小問（項目）からなるテストを考える。ここでは、各 A 点満点の項目が N 個集めたテストを考える。受験者 $i=1, \dots, N$ の項目 $j=1, \dots, M$ の点数を 0 以上の整数 X_{ij} とする。項目 j の特性値は次のように計算される。

$$\text{項目容易度} = \frac{1}{A} \sum_{i=1}^N X_{ij},$$

項目弁別力

$$= \frac{1}{A} \left\{ \sum_{(X_i \text{が上位}(\frac{N}{2}) \text{の} i)} X_{ij} - \sum_{(X_i \text{が下位}(\frac{N}{2}) \text{の} i)} X_{ij} \right\},$$

項目弁別係数 = X_{ij} と X_i の相関係数、

$$\text{Cronbach の} \alpha = \frac{M}{M-1} \left(1 - \frac{\sum_{j=1}^M \sigma_j^2}{\sigma^2} \right),$$

ただし、合計点 $X_i = \sum_{j=1}^M X_{ij}$ 、 $X_i (i = 1, \dots, N)$ の分散を σ^2 、 $X_{ij} (i = 1, \dots, N)$ の分散を σ_j^2 と書いた。

項目弁別力、項目弁別係数は、学力の差に対する点数の感性を表し、Cronbach の α はテスト全体の信頼性を表す。

3. システムの仕様

本システムは、解答パターン行列 X_{ij} と、各項目の満点 A_i を CSV 形式のファイルとして受け取る。

本システムは、Web ページ上に分析結果を表示する。特に次の特性値と図を含む。

- 項目の特性
 - 項目容易度、項目平均点、項目平均点、項目標準偏差、項目弁別力、項目弁別係数
 - 項目点数分布、項目特性曲線、項目と総得点の散布図
- テストの特性
 - テストの信頼性係数
 - テストの総得点分布
- 項目と受験者の特性
 - S-P 表

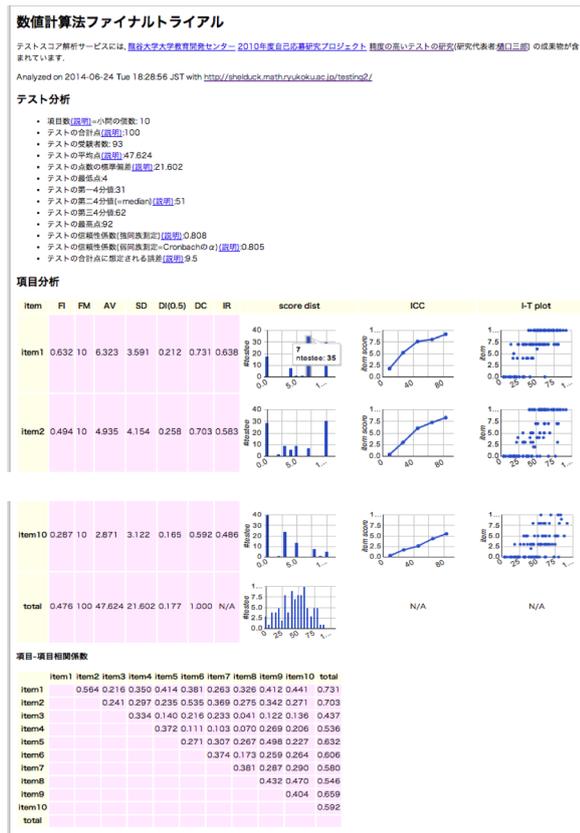


図 1 分析結果のスクリーンショット

4. システムの実装

本システムの開発段階での実装は文献⁽⁴⁾で報告した。Apache HTTP サーバ、MySQL データベース管理システム、サーバサイドスクリプト PHP で実装されている。グラフは Google Chart API を利用して SVG で Web ページに描画している。分析結果ページのスクリーンショットの一部を図 1 に示す。

4.1 HTML ファイルのポータビリティ

本システムが生成する分析結果は、教授者が利用する以外に、学習者やコース評価者、カリキュラム編成担当者に提供するという用途を想定している。ブラウザの機能で PDF 形式に変換して提供できることはもちろんであるが、付随するスタイルシートや画像ファイルを必要としない単一の HTML ファイルであるため、(JavaScript を含む)HTML コードをコピーして、LMS、CMS 等の HTML エディタに貼り付けることが可能であり、多くの用途に利用できる。

4.2 プライバシーに対する考慮

アップロードするデータは、学生者を特定する情報を含む必要がない。しかし、学習者数が少ないコースにおいては、他のデータとつぎあわせて個人に対する情報を抽出できる可能性がある。

このことから、アップロードしたデータは PHP のセッション変数としてセッションが持続する間のみ保持し、破棄するようにしている。また、分析結果に小問の点数-合計点の散布図を含む場合には、アップロードしたデータすべてを HTML ファイルが含むため、散布図の有無を分析者が選択できるようにしている。

4.3 分析者に対する考慮

分析者にテスト理論の予備知識を要求しないため、各特性値の説明と解釈のアドバイスを含んでいる。

5. システムの評価

本システムの評価は、現在のところ予備的な段階にとどまっている。龍谷大学理工学部の入学時数学プレースメントテストの結果の分析に、テスト理論の知見のない分析担当者が利用したところ、特別なインストラクションなしに使用することができ、仕様通りに動作することが確認された。また、分析担当者が各項目に対する知見を得たことが観察された。

6. おわりに

本システムは、別に用意したデータを分析できるという点では、R のパッケージや、項目応答理論の専用ソフトウェアと共通しているが、Web ブラウザさえあれば使用できる Web アプリケーションであって、少ない学習コストで、無償で使用できる点で優れている。一方、分析の範囲、柔軟性の点ではこれらに劣る。

本システムの拡張として、項目応答理論に基づく分析機能を付与することが考えられる。その際には、いくつか検討すべき点がある。項目応答理論ではデータの量が多く必要であるため、複数回の項目の使用にわたってプライバシーに配慮しつつ実施結果を蓄積する仕組みとする必要がある。また多くの計算が必要であるため、サーバ側ですべての計算する現在のシステムでは、同時利用者数に対する制限が厳しくなる。

本システムの開発の初期段階は、龍谷大学大学教育開発センター2010年度FD自己応募研究プロジェクト⁽⁴⁾で行われた。

参考文献

- (1) 池田央：“現代テスト理論”，行動計量学シリーズ，朝倉書店 (1994)
- (2) 豊田秀樹：“項目反応理論[入門編]”，朝倉書店 (2012)
- (3) 植野真臣，莊島宏二郎：“学習評価の新潮流”行動計量の科学 4，朝倉書店 (2010)
- (4) 樋口三郎：“精度の高いテストの研究”，FD・教材等研究開発報告書第 13 号，龍谷大学大学教育開発センター(2011)