

# 項目反応理論を用いたテスト結果の解析と妥当性の検証

小 玉 成 人\*・北 川 翔 太\*\*

## Development of WBT Teaching Material Using Item Response Theory.

Naruhito KODAMA\* and Shouta KITAKAWA\*\*

### Abstract

This paper presents development of WBT teaching material using item response theory. Recently, the difference of student's proficiency by a decrease in the number of students is growing. Therefore, the teaching material making that exists in each student's proficiency to be difficult to raise the learning effect in the same content of the lecture, and to raise the learning effect is needed. Then, because it aims at the improvement of the learning effect by making it study objectively by using the e-learning system and the item reaction theory around the field for a good field, and a constant effect was able to be achieved, it reports in this research.

**Keywords** : item response theory, e-learning, teaching material

### 1. まえがき

近年、少子化が問題になり、学生減少による各学生の習熟度の差が大きくなっている。そのため、同じ講義内容では学習効果を上げることが難しく、学習効果を上げるためには各学生の習熟度にあった教材作成が必要となる。しかし、学生毎に教材を作成しては教員の負担が増え結果として教育の質を落とすことになりかねない。そこで、教員が直接関わらずとも学生が自主的に学習できる e-learning を用いる方法が考えられる。現在、コンピュータやインターネット、モバイル端末などの情報通信技術 (ICT) を用いた ICT 活用教育を行っている教育機関は 75.8% にのぼり、これらの ICT を活用し、学

習者が主体的に学習できる e-learning を行っている教育機関は 51.1% と年々増加している。さらに、現行の制度では通学制の場合 124 単位中 60 単位まで e-learning を含むメディアを利用して単位認定を行うことができ、e-learning を行っている教育機関の中で、実際に単位認定を行っている教育機関は 40.4% になる<sup>(10)</sup>。しかし、e-learning はコンテンツを準備する労力が大きく、また個別の学生の習熟度に合った教材を用意することは難しい。そこで、本研究では、e-learning システムを用いて能力判定テストを行い、その結果から項目反応理論を用いて客観的に苦手分野を求め、その分野を中心に学習させることによって学習効果の向上を目指し、一定の効果があげることができたので報告する。

---

平成 21 年 12 月 14 日受理

\* システム情報工学科・講師

\*\* システム情報工学科・4 年

### 2. e-learning システムの概要

今回、e-learning システムの構築に LMS

(Learning Management System) でコース管理が容易にできる Moodle を採用した。図 1 は Moodle を用いるためのシステム構成である。Moodle は、利用者や管理者がネットワークを介してアクセスすることで、容易にサイトの利用や管理を行うことができ、インストール、アップグレードが容易であり、自動インストールパッケージがいくつか提供されている。また、無償で利用することが可能であり、多くのモジュールは追加費用なしにインストールすることができる。Moodle は、教材管理のほか、クイズ形式の問題作成機能や会議システム(フォーラム) など各種機能を持つ。

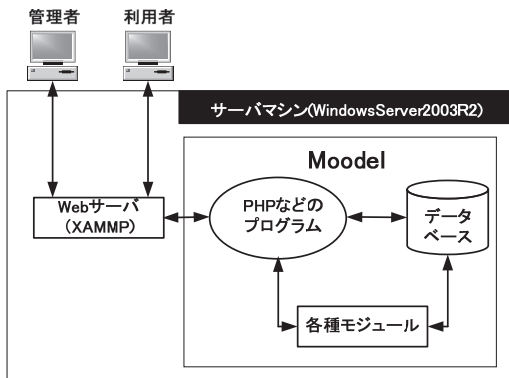


図 1 e-learning システムの構成

### 3. 実施した演習内容

本研究は、以下の開発環境で行った。

表 1 開発環境

	サーバ	管理者
OS	Windows Sever 2003 R2 SP2	Windows XP Home Edition Version2002 SP3
CPU	Intel(R) Pentium(R) 4 3.20GHz	Intel(R) Core(TM)2 Duo 2.00GHz
メモリ	2GB	2GB
HDD 容量	150GB	75GB

今回の実験では、能力判定テストを講義を受講しているおよそ 60 名を対象に一斉に行うため、サーバに対する負荷テストを行った。図 2 にサーバ負荷テスト時の使用可能なメモリの容量を示す。図に示されているように、使用可能メモリは 1GB を超えることは無く、2GB のメモリで十分な余裕があることが分かる。

また、図 3 は IP v 4 でのデータグラム (パケット) の毎秒あたりの送信量である。使用したネットワーク環境は、スイッチ、NIC とともに 1G bps で通信している。60 人が一斉にアクセスした結果、図のように最高で約 10,000 パケット (1 パケット 1024bps = 10,240,000bps) であった。したがって、ネットワーク環境は、60 人が同時にアクセスした場合にも問題ないことが分かった。

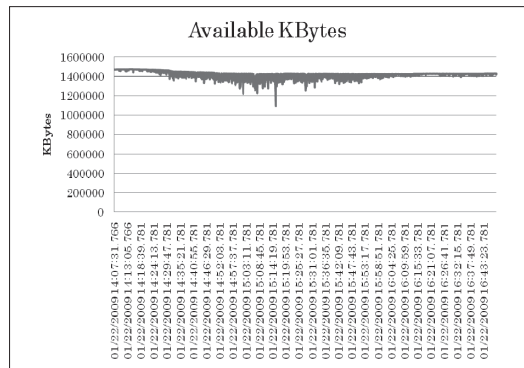


図 2 使用可能なメモリ容量

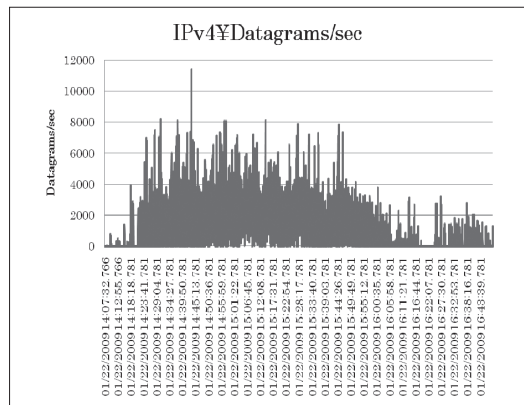


図 3 ネットワーク使用量

## 4. 項目反応理論の適用

項目反応理論（IRT:Item Response Theory）は、評価項目群の応答に基づいて、被験者の特性と評価項目の難易度を測定するためのテスト理論である。IRT は被験者やテストの内容に依存せず、不変的に被験者の能力とテスト項目の難易度を求められるため能力値の判定に利用した。

### 4.1 1パラメータロジスティックモデル － 1PL モデル

項目反応理論の数理モデルには、パラメータによって分けると 1PL（パラメータロジスティック）、2PL、3PL モデルがある。今回は、計算を単純化するため、最も単純で少ないサンプル数でも適切な計算ができる 1PL モデルを用いた。1PL モデルでは項目  $i$  の正解率は以下の式で表される。

$$P_i(\theta) = \frac{1}{1 + e^{(-Da(\theta - \beta_i))}} \quad (1)$$

ここで (1) 式の  $D$  と  $a$  は定数で  $D=1.7$ 、 $a=1$  である。また、 $\beta_i$  は項目難易度を表し、 $\theta$  は被験者能力を表している。従って、正解の確率は被験者能力  $\theta$  と項目難易度  $\beta_i$  の差（ $= \theta - \beta_i$ ）により決まり、被験者の能力が項目難易度より大きければ正解の確率は高くなり、逆に被験者の能力が項目難易度より小さければ正解の確率は低くなることを意味している。また、特に  $D=a=1$  の場合を Rasch モデルと呼んでいる。

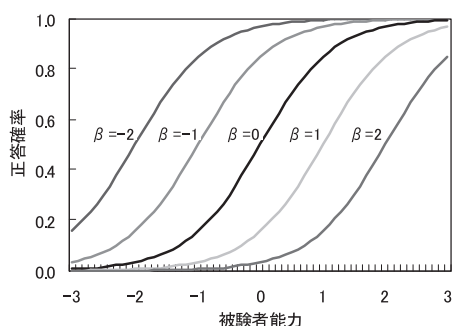


図4 1パラメータロジスティックモデル

## 4.2 採点方式

採点方式で分けると 2 値採点モデル、段階反応モデル、混合モデルなどがある。2 値採点モデルは、その項目が正答したら 1、誤答の場合 0 とする採点方式である。テスト項目は互いに独立であり、その項目だけで答えを出すことができるような問題に対応する。つぎに、段階反応モデルは、前の問題が次の問題にも影響するもので、段階数の増加に従って、答えのパターンも増加する。最後に、混合反応モデルは前述した 2 つのモデルを混合したモデルである。今回は、計算処理を簡潔にするため 2 値採点方式を採用した。

## 4.3 パラメータの推定

4.1 で述べたように数理モデルは 1PL モデルを用いることにした。なお、モデルのパラメータ推定方法には、様々な方法が開発されているが、本研究では簡単な手計算でも行うことができる PROX 法を用いた。

## 5. 項目反応理論を用いた学習方法の提案

### 5.1 学習の流れ

学習の流れは、図 5 に示すように始めに 1 回目の能力判定テストを行い、その結果から項目

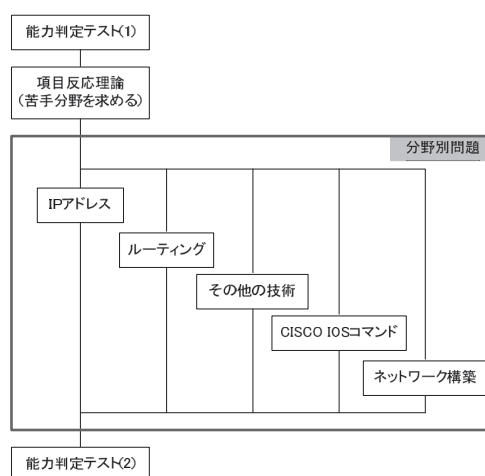


図5 学習フローチャート

表 2 学習分野

分野名	概 要
IP アドレス	IP アドレスに関する基礎的な知識や計算問題など
ルーティング	スタティックルーティングやダイナミックルーティングなどの基礎知識など
その他の技術	VLAN や NAT、DHCP などのルーティング以外のネットワーク技術など
CISCO IOS コマンド	CISCO ルータの IOS コマンドに関する問題など
ネットワーク構築	小規模ネットワークを構築するための機器の接続や IP アドレスの配分方法など

反応理論を用いて受験者の能力を計算する。受験者の能力は分野毎に計算し、その結果から学生は不偏的に判断された自分の能力値の最も低い分野を重点的に学習することができる。学習後には、2 回目の能力判定テストを行い本研究方法の効果を確認した。

なお、対象とした講義は、ネットワーク関連資格の CCNA の問題形式が 2 値採点モデルに似ているため CCNA 取得に関する講義を採用した。また、学習分野は表 2 に示すように 5 つの分野に分けた。

## 5.2 Moodle を用いた e-learning システムの構築

前述した Moodle を用いて作成した e-learning システムを図 6 に示す。この e-learning システムには各個人を識別するため

のログイン機能、講義の教材を閲覧するための機能、能力判定テストを行うための小テスト機能、問題をストックしておくためのデータベース機能などがある。

### 5.3 項目反応理論による能力判定

まず始めに、能力判定テストでの採点方法を 2 値に変更する必要がある。Moodle の小テスト機能は、一回間違えると -0.1 され、正解すると +1.0 となる。そのため、不正解の場合は「0」、正解の場合は「1」となるが、テスト中に何回でも回答することができるため、0.9 以下は 0 と判断する。また、2 回以上同じテストを行っている学生もいるが、今回は最低点を用いるものとする。

つぎに、能力判定テスト（1）を行い、以下で定義されるロジット・インコレクト、ロジット・コレクトを用いて線形化し、分野別に学生の能力値を求めた。

$$\text{ロジット・インコレクト} = \ln((1-p)/p) \quad (2)$$

$$\text{ロジット・コレクト} = \ln(p/(1-p)) \quad (3)$$

ここで、 $p$  は正答率を意味する。

最後に、能力値の低い分野を集中的に勉強してもらった後、再度能力判定テスト（2）を行い、同様の計算方法で能力値を求めた。能力判定の結果を表 3 に示す。表から 1 回目の最終能力の平均値は 0.164 に対し、2 回目は 0.271 と上昇し、本学習方法の効果が示された。しかし、今回は講義時間の都合上、1 分野につき 6 問で能



図 6 作成した講義の e-learning システム

力判定を行っているため、より正確な能力判定を行うためには問題数を増やす必要があると考えられる。また、本来の受講学生は60名程度だったにもかかわらず自主的に学習してもらったため、途中で能力判定テストを止めたり、1回目か2回目の試験を行っていなかったりと有効な結果が得られた学生が20名程度となってしまった。そのため、ある程度の強制力を持たせる必要があると思われる。また、今回はサンプル数の関係でIRTの1PLで計算したが3PLなどの高機能なモデルを用いた方法も考慮する必要がある。

表3 最終能力の比較

分野名	最終能力（1）	最終能力（2）
IP アドレス	0.461	0.868
ルーティング	-1.213	-0.117
その他の技術	-0.175	0.805
CISCO IOS コマンド	0.116	0.278
ネットワーク構築	0.538	0.271
全体	0.164	0.271

## 6. テスト結果の妥当性検証

項目反応理論で計算を行う場合、幾つかの前提条件がある。項目反応理論の前提条件は以下の2つである。

- ①局所独立の仮定
- ②一次元性の仮定

これらは、「ある問題に正解できる確率は、他の問題に正解できる確率の影響を受けない」と「全ての問題はただ一つの能力分野を測定するものである」とことである。これらに関しては、一問ずつ独立した問題を作成しているため問題無いと思われる。

つぎに、モデルとして利用した1PLモデル(Rasch Model)の前提条件は以下になる。

- ①問題の弁別力は同一
- ②当て推量が最小
- ③学生が多い

①は、問題によってロジスティック曲線の形が変わらないということである。いくつかの問題をピックアップした結果、図7に示すようなロジスティック曲線となったため、問題ないと思われる。②については、偶然正解する確率が無いことを意味しているが、今回作成した問題は4択問題であるため25%の確率で偶然正解することになる。そのため、当て推量を考慮した2PLモデルなどの利用を考慮する必要がある。③については、標本数が100～200程度必要となっており、今回の標本数では少し少ないことが分かるが、本学では1クラス70名程度なのでこれ以上増やすことは難しく、無効なデータをできるだけ減らすことやデータが足りないことでの影響の度合いを調査する必要があると思われる。

最後に、計算に用いたPROX法的前提条件として以下の2つがある。

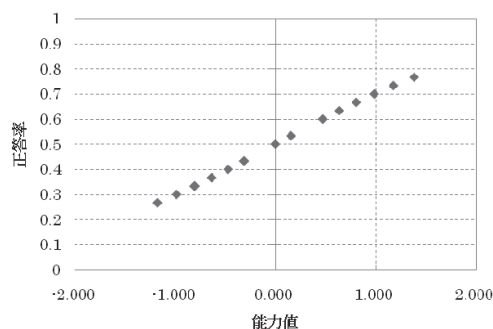


図7 正答率と能力値の関係

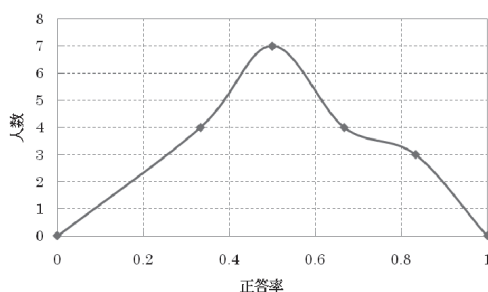


図8 能力値の分布



①問題の困難度が正規分布

②能力が正規分布

図 8 は IP アドレス分野の能力値の分布である。図よりおおよそ正規分布になっていることが分かる。

## 7. まとめ

学生の多様化に対応し、学習効果を上げるため、被験者やテストの内容に依存しない項目反応理論を用いて e-learning システムを構築した。この教材を用いて学習してもらった結果、最終能力は平均 0.107 向上した。しかし、e-learning システム上で分野別の問題を準備し、自主勉強してもらったため、分野別問題を行わなかった被験者が多く、多くのサンプル数を得られなかった。そこで、自主学习ではなくブレンディッド型の授業を行うなどにより、ある程度強制力を持たせる必要があると思われる。また、今回は能力判定テスト後、手計算で苦手分野を求めたために、完全な形で e-learning システムを構築できなかった。そこで今後は、自動的に最終能力を求める Moodle モジュールとその結果から問題を生成するモジュールの作成を行う予定である。

## 参考文献

- (1) 尾崎, 松坂:「項目反応理論による数学の基礎能力の推移分析」, 八戸工業大学紀要, 第 27 巻, pp.61-67, (2008-8)
- (2) 佐々木:「項目反応理論による e-learning コンテンツの配信制御」, 情報処理学会東北支部第 7 回研究会資料, No.10, (2007-12)
- (3) 吉田:「項目反応理論を用いた WBT 教材の作成」, 八戸工業大学工学部システム情報工学科卒業研修論文, (2009-2)
- (4) 尾崎, 松坂:「項目反応理論による数学の基礎能力の推移分析」, 情報処理学会東北支部第 7 回研究会資料, No.9, (2007-12)
- (5) Naruhito Kodama, Tomoyuki Matsuzaka, Takayuki Iqanuma, Noburu Kurihara, Yasuhiro Ozaki:「Online Education for Students and Community People Using an E-Learning System」, Proc. of ITHET2007, Kumamoto, (2007-7)
- (6) Takayuki Iwanuma, Tomoyuki Matsuzaka:「Web Based Education Method on Wind Energy Using an E-Learnig System」, Proc. of Renewable Energy 2006, October, Makuhari, (2006.10)
- (7) 大友:「言語テスト・データの新しい分析法 項目反応理論入門」, 大修館書店, (1996)
- (8) Frank B. Baker, Seock-Ho Kim:「Item Response Theory: Parameter Estimation Techniques」, Marcel Dekker Inc, (2004)
- (9) 豊田:「項目反応理論[理論編]」, 朝日書店, (2005)
- (10) 日本イーラーニングコンソーシアム:「eラーニング白書 2008/2009 年版」, 東京電機大学出版局, (2008)
- (11) 井上, 奥村, 中田:「Moodle 入門」, 海文堂, (2006)