

# 項目反応理論による数学の基礎能力の推移分析

尾崎 康弘\*・松坂 知行\*

## Time Trace Analysis of Basic Mathematical Ability Using Item Response Theory

Yasuhiro OZAKI\* and Tomoyuki MATSUZAKA\*

### Abstract

With the rapidly declining population of students in the 18-year-old bracket, the applicants to the university are decreasing, and as a result, the students who do not have adequate mathematical ability to learn in colleges are increasing, and thus teaching staffs are forced to open remedial classes for them. This paper treats a time trace analysis of basic mathematical ability of the students with item response theory, considering year-on-year comparisons.

**Keywords:** Item response theory, Mathematical ability, Time trace analysis

### 1. はじめに

18歳人口の減少に加え、大学の理工系学部への志願者が減少し、理工系の基礎を学ぶために必要な数学の知識が、必ずしも十分ではない学生が大学に入学するようになってきた。そこで、大学教育を受けるために最低限必要な知識を身につけさせるため、リメディアル教育が必要になり、これを実施する大学が増えてきている。

本学では、昭和50年以来、基礎学力を把握するため、1学年を対象として、入学直後に、数学、物理、化学、国語、英語の開講試験を行ってきた。開講試験の結果は、習熟度別クラスの編成、学生の能力に見合った独自のテキストの作成、教育方法の改善などに活用されてきた。開講試験の内容は、実施初年度以来、基本的に変えていないので、学生の学力の推移を検討するデータとして有用であると考えられる。

そこで、今回数学を対象科目としてとりあげ、

過去10年間に亘る入学者の能力の推移を調査することにした。学生の能力の評価方法には、情報処理技術者試験、TOEIC、TOEFLなどの評価に使われている項目反応理論を用いることにした。項目反応理論を用いることにより、テスト結果から、問題毎の難易度と被験者の能力を分離して推定できるので、入学者の能力の年度による推移を調査する方法として適切であると考えた。

本稿は、本学平成8年から平成19年までの開講試験で行われた数学の成績を、項目反応理論を用いて分析し、その結果から数学能力を育成するための対応策について述べた。

### 2. 開講試験の内容

開講試験は、数学に関する学生の基礎学力を把握するために考案され、実施されてきた。対象学生は、全入学生である。この試験内容は、学生の基礎学力状況を把握するために必要な基本的な内容を中心に構成されている。この概要を表1に示す。

---

平成19年12月14日受理

\* システム情報工学科・教授

表 1 開講試験の問題

項目	内容
1	比例式の計算
2	絶対値の計算
3	指数（整数）の計算
4	指数（小数）の計算
5	指数（分数）の計算
6	常用対数の計算
7	非常用対数の計算
8	度表示の三角関数の計算
9	ラジアン表示の三角関数の計算
10	2 次式の因数分解
11	連立 1 次方程式の求解
12	2 次関数の頂点の求解
13	2 次関数の y 軸交点の求解
14	2 次関数のグラフ
15	直交する 2 直線の勾配の求解
16	直交する 2 直線の切片の求解
17	微分の計算
18	不定積分の計算
19	極限値の計算
20	部分積分の計算

この開講試験は、講義開始前に全学的な行事として行われ、この結果をもとにして、本人の希望をとり入れながら学科横断的なクラス編成を行い、教育効果を上げるように配慮している。全学科の学生を対象としているため人数が多く、迅速な結果を得るため、成績処理には、マークカードリーダーやマークシートリーダーを用いた。

この試験結果は、入学生の学力分布を把握するのに有効であった。特に、学力下位の学生に対する学力状況を把握する上で大いに役立った。さらに、この試験結果は、クラス編成だけでなく、多様な学生を教育指導するための独自

の教科書・演習書の作成や講義の工夫など教員側の指導方法にも生かされてきた。

### 3. 項目反応理論の概要<sup>(1),(2),(3)</sup>

本章では項目反応理論における項目特性曲線、パラメータ推定法、テスト情報関数について述べる。

#### 3.1 項目特性曲線

項目反応理論（略称 IRT=Item Response Theory）は、テストでの評価項目群への応答に基づいて、被験者の能力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 は被験者の母集団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という利点がある。IRT では被験者の正答確率は (1) 式のように与えられる。

$$p(\theta) = \frac{1}{1 + e^{-(\theta - \beta)}} \quad (1)$$

$p(\theta)$ : 正答確率  $\theta$ : 被験者能力  $\beta$ : 難易度

図 1 は (1) 式の正答確率を図示したもので、被験者能力と問題の難易度が等しいときには 0.5 になり、被験者能力の方が難易度よりも高い場合は高くなり、逆に被験者能力の方が難易度よりも低い場合は低くなることを示している。つまり正答確率は被験者能力と問題の難易度の差によって決まることを表している。このモデルは最も簡単なモデルで、Rasch モデルと呼ばれている。

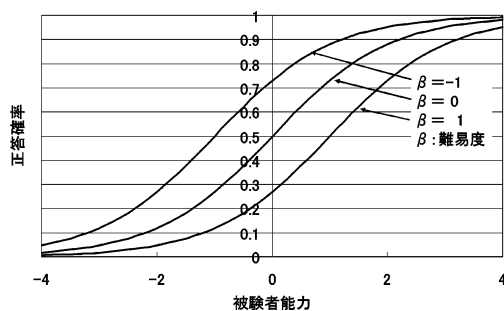


図 1 項目特性曲線

表2 試験項目と被験者の2値採点

	試験項目				
被験者	1	2	3	...	$n$
1	1	0	1	...	1
2	0	1	1	...	0
3	1	1	0	...	1
⋮	⋮	⋮	⋮	...	⋮
⋮	⋮	⋮	⋮	...	⋮
⋮	⋮	⋮	⋮	...	⋮
$N$	1	0	0	...	1

いま被験者の数を  $N$ 、試験項目の数を  $n$  とすると (1) 式は (2) 式のように変形できる。さらに、IRT では、試験項目の得点は、正解が1、誤答が0 で与えられ(2 値採点)、(2) 式の  $u_{ij}$  は、このことを表す変数である。表2 は試験項目と被験者の2 値採点の様子を示す。

$$P(u_{ij} | \theta_j, \beta_i) = \frac{\exp[u_{ij}(\theta_j - \beta_i)]}{1 + e^{(\theta_j - \beta_i)}} \quad (2)$$

$\theta_j$  : 被験者能力 ( $j=1, 2, \dots, N$ )

$\beta_i$  : 項目難易度 ( $i=1, 2, \dots, n$ )

$u_{ij}$ : 正答のとき1, 誤答の場合0

### 3.2 被験者能力と難易度の推定

(2) 式を用いて正答確率を求めると、テスト結果は  $n \times N$  行列から成る反応パターンから成り、そのときの尤度関数は以下のようになる。

$$l = \prod_{j=1}^N \prod_{i=1}^n p(u_{ij} | \theta_j, \beta_i) = \frac{\exp[\sum_{j=1}^N \sum_{i=1}^n u_{ij}(\theta_j - \beta_i)]}{\prod_{j=1}^N \prod_{i=1}^n [1 + e^{(\theta_j - \beta_i)}]} \quad (3)$$

しかし、(3) 式のままでは被験者能力と難易度の推定が難しいので、(3) 式の対数を取り対数尤度関数を求める。

$$L = \log l = \sum_{j=1}^N \sum_{i=1}^n u_{ij}(\theta_j - \beta_i)$$

表3 パラメータの推定方法

Maximum Likelihood Estimation	項目パラメータが既知のとき、それを用いて被験者能力を最尤推定する方法
	被験者能力が既知のとき、それを用いて項目パラメータを最尤推定する方法
Joint Maximum Likelihood Estimation	項目パラメータと被験者能力を同時に最尤推定する方法
Marginal Maximum Likelihood Estimation	被験者能力の周辺分布を利用して、項目パラメータを最尤推定する方法
Joint and Marginal Bayesian Estimation	同時および周辺分布によるベイズ推定法
Heuristic Method	ある仮定を必要とするが、計算の速い簡便法

$$\begin{aligned} & - \sum_{j=1}^N \sum_{i=1}^n \log[1 + e^{(\theta_j - \beta_i)}] \\ & = \sum_{j=1}^N \sum_{i=1}^n u_{ij} \theta_j - \sum_{j=1}^N \sum_{i=1}^n u_{ij} \beta_i \\ & - \sum_{j=1}^N \sum_{i=1}^n \log[1 + e^{(\theta_j - \beta_i)}] \end{aligned} \quad (4)$$

$u_{ij}$  は反応パターンから与えられるので、これを (4) 式に代入すると、 $n \times N$  行列から成る非線形方程式が得られるので、被験者能力  $\theta_j$ 、難易度  $\beta_i$  は Newton-Raphson 法を用いて求めることができる<sup>(3)</sup>。

$$\begin{bmatrix} (\hat{\theta}_g) \\ (\hat{\beta}_i) \end{bmatrix}_{t+1} = \begin{bmatrix} (\hat{\theta}_g) \\ (\hat{\beta}_i) \end{bmatrix}_t - \begin{bmatrix} [\hat{L}_{gg}] & [0] \\ [0] & [\hat{L}_{ii}] \end{bmatrix}_t^{-1} \times \begin{bmatrix} (\hat{L}_g) \\ (\hat{L}_i) \end{bmatrix}_t \quad (5)$$

ここで (5) 式の記号は以下の通りである。

$$\hat{L}_g = \frac{\partial L}{\partial \theta_g}, \quad \hat{L}_i = \frac{\partial L}{\partial \beta_i}, \quad \hat{L}_{gg} = \frac{\partial^2 L}{\partial \theta_g^2}, \quad \hat{L}_{ii} = \frac{\partial^2 L}{\partial \beta_i^2} \quad (6)$$

パラメータの推定方法には、表3 のようにいろいろな方法が提案されてきた。本稿で用いた方法は、項目パラメータ (Rasch モデルでは項目難易度) と被験者能力を同時に最尤推定する Joint Maximum Likelihood Estimation 法 (JMLE 法) を用いた。

表 4 項目得点と被験者グループの得点

		被験者グループの得点						行合計
		1	2	...	$g$	...	$n-1$	
項目 得点	1	$f_{11}$	$f_{12}$	...	$f_{1g}$	...	$f_{1,n-1}$	$f_{1.}=s_1$
	2	$f_{21}$	$f_{22}$	...	$f_{2g}$	...	$f_{2,n-1}$	$f_{2.}=s_2$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$i$	$f_{i1}$	$f_{i2}$	...	$f_{ig}$	...	$f_{i,n-1}$	$f_{i.}=s_i$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	$n$	$f_{n1}$	$f_{n2}$	...	$f_{ng}$	...	$f_{n,n-1}$	$f_{n.}=s_n$
列合計		$f_{.1}$	$f_{.2}$	...	$f_{.g}$	...	$f_{.n-1}$	

### 3.3 被験者能力と難易度の計算手順

被験者能力と難易度の計算手順は以下の通りであり、表 4 はこの手順で得られる結果である。

- (1) 各被験者の得点を求め、得点の合計点によりグループ 1, 2, ...,  $n-1$  に分類する。
- (2) 同一グループに属する被験者  $j$  について、各項目  $i$  ごとの部分合計得点  $f_{ij}$  を計算する。
- (3) 各グループの列合計  $f_{.1}, f_{.2}, \dots, f_{.n-1}$  を計算する。
- (4) 各項目毎の行合計  $f_{1.}, f_{2.}, \dots, f_{n.}$  を計算する。

JMLE 法のプログラムの作成には、文献 (3) を参照し、MATLAB<sup>(4)</sup> を用いて開発した。MATLAB を用いた理由は、この言語が行列の操作に優れており、また、Excel とのデータの互換性があることである。被験者の素点成績データは Excel シートの形で与えられるので、MATLAB との相性がよい。

### 3.4 テスト情報関数

Rasch モデルにおけるテスト情報関数は以下のように定義される。

$$I(\theta) = \sum_{i=1}^n P_i(\theta)(1 - P_i(\theta)) \quad (7)$$

また、(7) 式の各項目は項目情報関数と呼ばれ、以下のように定義される。

$$I_i(\theta) = P_i(\theta)(1 - P_i(\theta)) \quad (8)$$

テスト情報関数は、一般的な統計学の分野でフィッシャー情報量と呼ばれているものと同じである。テスト情報関数は、値が大きければ大きいほど信頼性のあるテストであることを示している。また、テスト情報関数を用いると、推定の標準誤差  $\delta(\theta)$  が以下の式で与えられる。

$$\delta(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (9)$$

## 4. 分析結果

以下に、項目反応理論による分析結果を述べる。

### 4.1 採点方法

すでに述べたように IRT では正解が 1, 誤答が 0 という 2 値採点で行う。本稿では、平成 19 年度、15 年度、11 年度の 4 年間隔の開講試験および平成 11 年度から 3 年遡った平成 8 年度(これは現在保存している最も古いデータである)の開講試験を対象とした。各年度の試験問題は全く同一であり、また、問題数は 20 問である。各問題は独立に回答でき、正解のとき 1, 誤答のとき 0 の得点を与えた。対象とした学科は、改組などによる組織改変のない 3 学科を選んだ。

### 4.2 得点のグループ化

まず、得点により被験者をグループ化する。テスト問題は 20 問であるため、0 から 20 グループに分類できるが、最尤推定法では全問満点の 20 グループと全問零点の 0 グループを除外しないと被験者能力と難易度の推定ができない。そこでグループ数を 1 から 19 グループに分類し、グループ毎の被験者能力を求めることとした。

### 4.3 被験者能力

つぎに、各グループに属する被験者数に対する全被験者数の割合を求め、この相対値を比較することによって年度毎の能力の推移を表現することにした。

図2は、横軸に被験者能力、縦軸にはその被験者能力に属する被験者数の割合を年度ごとに示したものである。平成8年度には、かなりの学生は能力零以上のグループに属するが、平成11年度にはこの数は半数程度になり、平成15年度以降には、能力零以下のグループに属する学生が非常に増えていることが分かる。

図3は、被験者能力零未満の割合の推移を示したグラフであり、平成15年度からは学生の約70%が被験者能力零未満に属し、急速に数学力が下がっていることを示している。

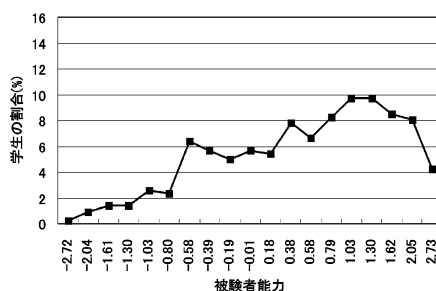
### 4.4 項目難易度

図4は、各試験項目に対する難易度を求めたものである。年度により値が若干変わるが、傾向は全く同じなので平成19年度のデータで求めたものを示した。通常、難易度は $-1.4 \sim +1.4$ の範囲に分布すると言われており $-1.4$ は易しい問題、 $+1.4$ は難しい問題とされている。図4から判断すると、極端にやさしい問題や難しい問題はないことが分かり、開講試験は、適切なレベルに分布していると思われる。

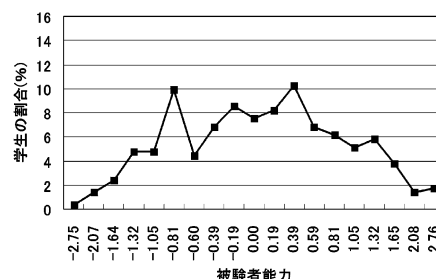
表1を参照すると、指数(小数, 分数), 対数, 三角関数が難易度0.5以上であり、やや難しい問題である。部分積分は1.0以上であり、難しい問題と思われる。また、比例式の計算, 指数(整数), 2次式の因数分解, 連立1次方程式の求解は、 $-0.5$ 以下であり易しい問題と受け止められている。

### 4.5 被験者グループの正答確率

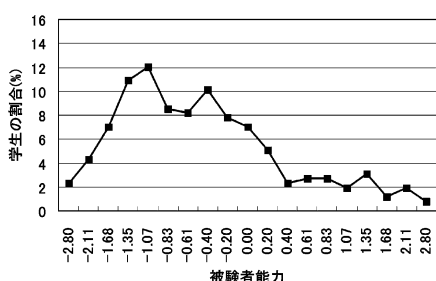
被験者能力と項目難易度が求められたので(1)式から、各被験者グループの各項目に対する正答確率を計算できる。表5は被験者グルー



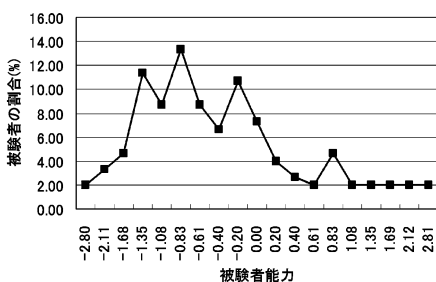
(a) 被験者能力と成績の割合（平成8年）



(b) 被験者能力と成績の割合（平成11年）



(c) 被験者能力と成績の割合（平成15年）



(d) 被験者能力と成績の割合（平成19年）

図2 被験者能力と被験者の割合

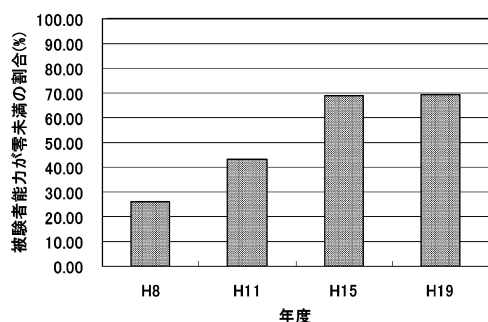


図3 被験者能力零未満の割合の推移

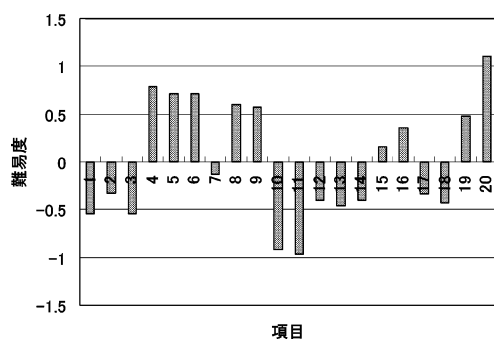


図4 各項目に対する難易度

表5 被験者グループの正答確率

		被験者のグループ別能力				
		1	2	3	...	19
項目	難易度	-2.8030	-2.1149	-1.6827	...	2.8112
1	-0.5476	0.0949	0.1726	0.2432	...	0.9664
2	-0.3236	0.0773	0.1429	0.2044	...	0.9583
3	-0.5476	0.0949	0.1726	0.2432	...	0.9664
4	0.7858	0.0269	0.0521	0.0781	...	0.8834
5	0.7092	0.0290	0.0560	0.0838	...	0.8911
.	.	.	.	.	...	.
.	.	.	.	.	...	.
.	.	.	.	.	...	.
20	1.1054	0.0197	0.0384	0.0580	...	0.8463

プ1, 2, 3, ..., 19の各項目に対する正答確率を算出したものの一部である。この表から分かるように、被験者グループに属する個々の学生が既に分かっているため、個々の学生の正答確率を把握でき、個別指導に役立てることができる。

#### 4.6 項目情報関数

項目情報関数は(6)式を用いて求められる。図5は、参考までに項目4, 11, 15, 20に対する項目情報関数を示したものである。ピーク値は、それぞれの項目難易度のところで現われている。

#### 4.7 テスト情報関数

つぎに、開講試験全体のレベルの妥当性を検討するため、テスト情報関数を求めた。図6はテスト情報関数を示す。最大値は、被験者能力が零付近で現われ、したがって被験者能力零付近で推定された被験者能力の値は、情報量が大

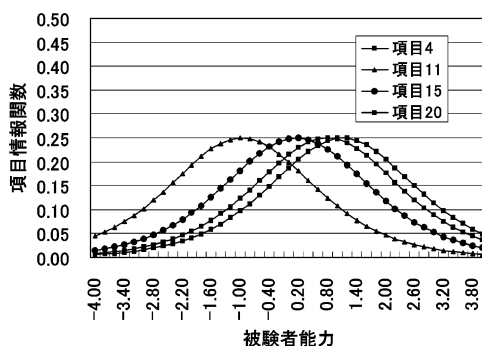


図5 項目情報関数

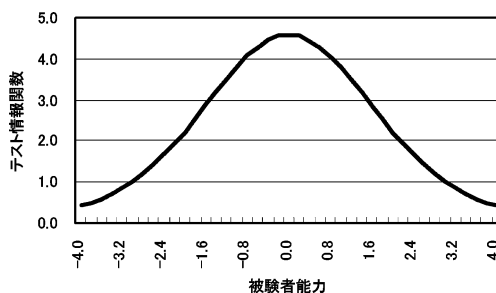


図6 テスト情報関数

きく信頼性があると判断できる。

また、被験者能力零を中心に、左右対称になっているので、試験のレベルは妥当な内容と思われる。最大値は約 4.5 であるので、標準誤差は 0.47 となり、被験者能力零と判定された場合、真の能力は  $-0.47 \sim 0.47$  である可能性が高いことを示している。このテスト情報関数は、平成 19 年度のデータで求めたものである、他の年度でもほぼ同様な結果であった。

## 5. ま と め

以上、項目反応理論を用いて数学能力の推移を求めた結果をまとめると

(1) 項目反応理論により被験者能力と項目難易度をはっきり分離することができた。その結果、被験者能力は、平成 15 年度頃から急速に低い方に分布しはじめ、平成 15 年度以降は被験者能力零以下の者が約 70% を占めていることが分かった。ちなみに、この数字は、平成 8 年度では約 27% であるから被験者能力零以下の者は約 2.6 倍に増加したことが分かった。

(2) 被験者能力と項目難易度から各被験者グループの正答確率表を求めることができるので、各項目ごとの正答確率の低い被験者を抽出し、個別的な対応が可能になると思われる。

(3) 従来、古典的な統計的テスト理論のもとに様々な指標で成績を評価していたが<sup>(5)</sup>、項目反応理論を用いると、被験者能力、項目難易度という単純かつ汎用的な指標でテスト結果を評価できるようになった。

(4) この手法をさらに発展させれば、より細かい項目で詳細な被験者能力の識別が可能であ

る。また、より効率の良い教育方法の策定も可能になる。たとえば、e-learning<sup>(6),(7),(8)</sup> を用いる教育方法もひとつであり、この際、正答確率に基づいて適応的に課題を提示するようにすれば個別指導に近い方法で教育指導できよう。

最後に、本学では、学生のために独自教科書や演習書を作成し、教育指導してきた。また、リメディアル教育（選択 2 単位の「数学基礎 I」・「数学基礎 II」）を新たに開設し、ナイトスクールなどの制度も導入し学生への教育指導対策を強化してきたが、将来的には、専従スタッフによる個別的指導を導入する必要がある。

## 参 考 文 献

- (1) 大友：項目反応理論入門，大修館書店，1996
- (2) 豊田：項目反応理論，朝倉書店，2006
- (3) Frank B. Baker, Seock-Ho Kim : Item Response Theory, Mercel Dekker, Inc., 2004
- (4) MATLAB：サイバーネット社 Web ページ
- (5) 尾崎康弘：授業方法の評価に関する一考察，東北数学教育学会年報第 31 号，pp. 52-56, 2000. 3
- (6) 尾崎康弘，松坂知行，高橋史朗：教育へのオンデマンド導入とコンテンツ，八戸工業大学紀要第 24 巻，pp. 275-281, 2005.2
- (7) Naruhito Kodama, Tomoyuki Matsuzaka, Takayuki Iwanuma, Nobuo Kurihara, Yasuhiro Ozaki : Online Education for Students and Community People Using an E-Learning System, Proc. of ITHET2007, Kumamoto, 2007, JAPAN, 2007.7
- (8) Takayuki Iwanuma, Tomoyuki Matsuzaka : Web Based Education Method on Wind Energy Using an E-Learnig System, Proc. of Renewable Energy 2006, October, 2006.10