

〔論 文〕

# 項目反応理論による習熟度別クラス分けの考察

中嶋 康博<sup>\*1</sup>・原 幸範<sup>\*2</sup>・境 優一<sup>\*3</sup>  
小川原弘士<sup>\*4</sup>・金井 政宏<sup>\*1</sup>

## Analysis of a Placement Test Under Item Response Theory

Yasuhiro NAKASHIMA<sup>\*1</sup>, Yukinori HARA<sup>\*2</sup>, Yuichi SAKAI<sup>\*3</sup>  
Hiroshi OGAWARA<sup>\*4</sup>, Masahiro KANAI<sup>\*1</sup>

### Abstract

Many students may experience difficulty in studying mathematics; however, engineering requires students to use mathematics. We expect the students to remediate basic mathematical ability. One of the authors has believed that there must be three groups of students under their motivations to learn, that effect directly on their attitudes. Hence we would have tried to classify them by the groups. In 2017, we actually introduced a placement test in our mathematics course in order to classify the students into three classes. Based on the results of our two examinations under item response theory, we determined the necessity of this classification. This article contains our opinions with respect to the classification.

**Key Words** : Mathematical Education, First Year Education, Remedial Education, Placement Test, Item Response Theory

## 1. 緒 言

2006年実施のPISA調査の結果から、生徒の科学的認識や科学への興味関心などについて課題があることが明らかになった。2008年の中央教育審議会答申では、理数科等の授業時数を増やすことや指導内容の充実が提言され、科学技術・理数教育充実のための取組が推進されている。そして翌2009年の学習指導要領改訂で理数科教育の充実が明示された。資源の乏しい日本において、次代を担う科学技術系の人材育成は不可欠であり、そのような期待を受けて工業大学においては社会で求められる工学専攻の学生の養成を積極的に進める必要がある。

一方、基礎的な数学の事項を修得していない学生が本学には少なからずおり、たとえば分数や平方根の計算ができない、解の公式を知らない、指数・対数やラジアン単位の意味を知らない、ということが講義中に生じることがある。しかし学生に尋ねると「高校卒業まで特に知らなくても問題なかった」、「高校ではラジアン単位なんて教えてもらえなかった」、そして「そんなことを知らなくても高校を卒業できたし、大学にも入学できた。だから問題ない」といった声が、残念ながら挙がってしまう。しかし工学を専攻する学生にとって、数学や物理の知識は必須である。なにかんずく学生がなすべきことを、数学や物理の基礎の基礎を確実に理解することだと考えるのは、ごく自然なことであろう。そのために学生の現状を把握し分析を行い、数学の基本能力をどのように育成すべきかを考察し、今後の教育につなげたいと考える。

## 2. クラス分け試験と定期試験

上述のような目的のもと、著者らは初年次学生の数学教育においてクラス分け試験を実施し、その結果を用いて習熟度別のクラス分けを行った。前期の講義をそのクラス分けに従って行い、定期試験の結果を得た。クラス分けテストおよび定期試験の結果を本稿で比較し、クラス分けの効果を考察する。

<sup>\*1</sup> 教育創造工学科, <sup>\*2</sup> 共通教育科, <sup>\*3</sup> 九州大学多重ゼータ研究センター, <sup>\*4</sup> 熊本大学大学院先端科学研究部  
平成29年12月6日受理

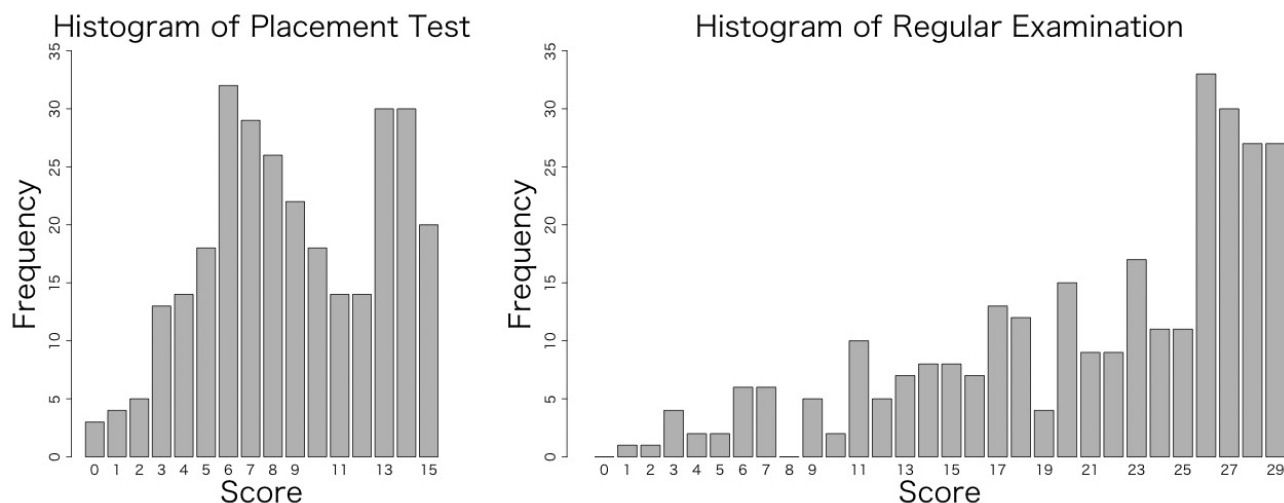


Fig. 1. Histograms of Placement Test and Regular Examination

その際に、著者らの興味がある項目反応理論 (Item Response Theory, IRT) を利用した。2016年度に本学で行った試験の処理<sup>(1)</sup>と同様の処理にて試験の各項目を評価したが、今回は受験者の特性値も評価の対象とした。単純な総得点ではなく IRT による特性値を利用した意図として、総得点では重複する受験者が複数おり差異の判断がしづらいうえ、特性値の方が『どの問題に正答したか』を表すパターン行列を基準にしており、難しい問題に正答したから高い値をとるなど、受験者の実態をより反映しようと考えた点がある。

さて、以下では2つの試験の概要を記述する。ともに2017年度の1年次の開講される講義で実施した。1つは4月に行ったクラス分け試験であり、受験者は本学4学科の319名であった。他の1つはクラス分け試験で習熟度別に分けた講義における定期試験であり、8月に行い受験者は316名であった。このうち、欠席者や再履修者を除いて、2つの試験をともに受験したものに制限すると292名になり、これが本稿で利用したデータである。クラス分け試験は15問からなり、定期試験は講義で扱う範囲の内容の29問 (1.1から4.4まで) からなる。いずれも基礎的な数学の内容で、すべての問題を小問としており、ある項目での誤答が別の項目の誤答を誘発することはない。可能であれば今後、各項目ごとに相関を見て、受験者の特徴的な傾向を調査できれば望ましい、と考える。IRT を利用するために、いずれの項目も正答を1、誤答を0として評価した。

図1は2つの試験のヒストグラムである。クラス分け試験に関しては、満点は15点であり、得点の平均は8.9、標準偏差は3.9であった。分布には山が2つあり、上限を15点にしているためその付近は右の裾が切れているものの、ど

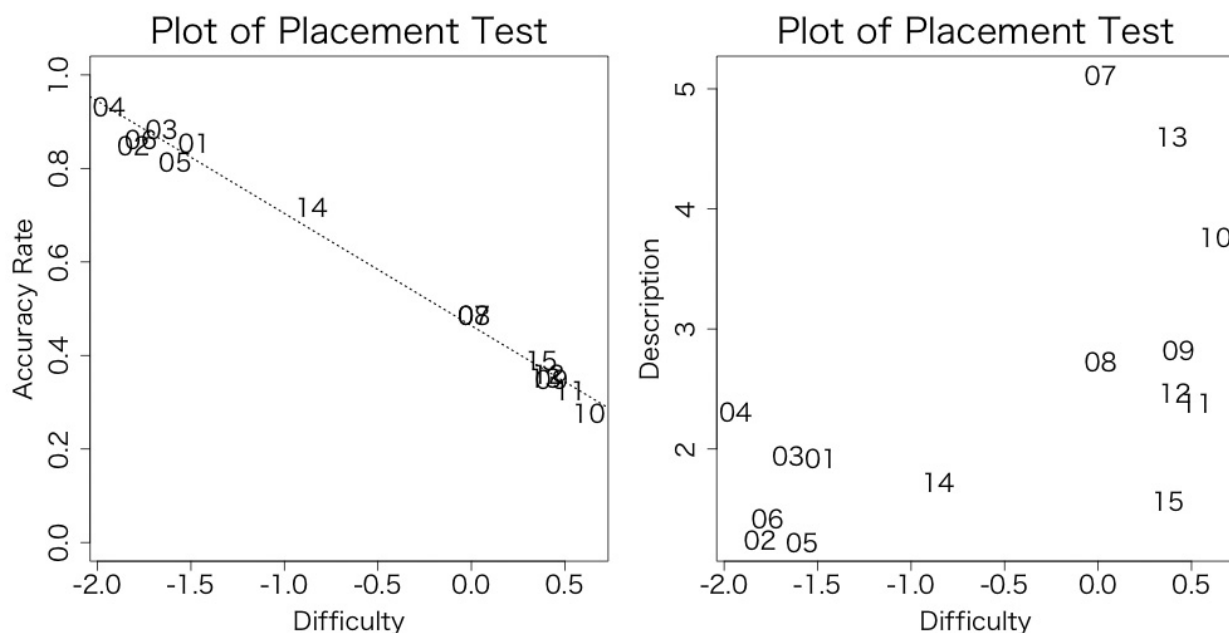


Fig. 2. Scatter Plots of Parameters of Placement Test under IRT

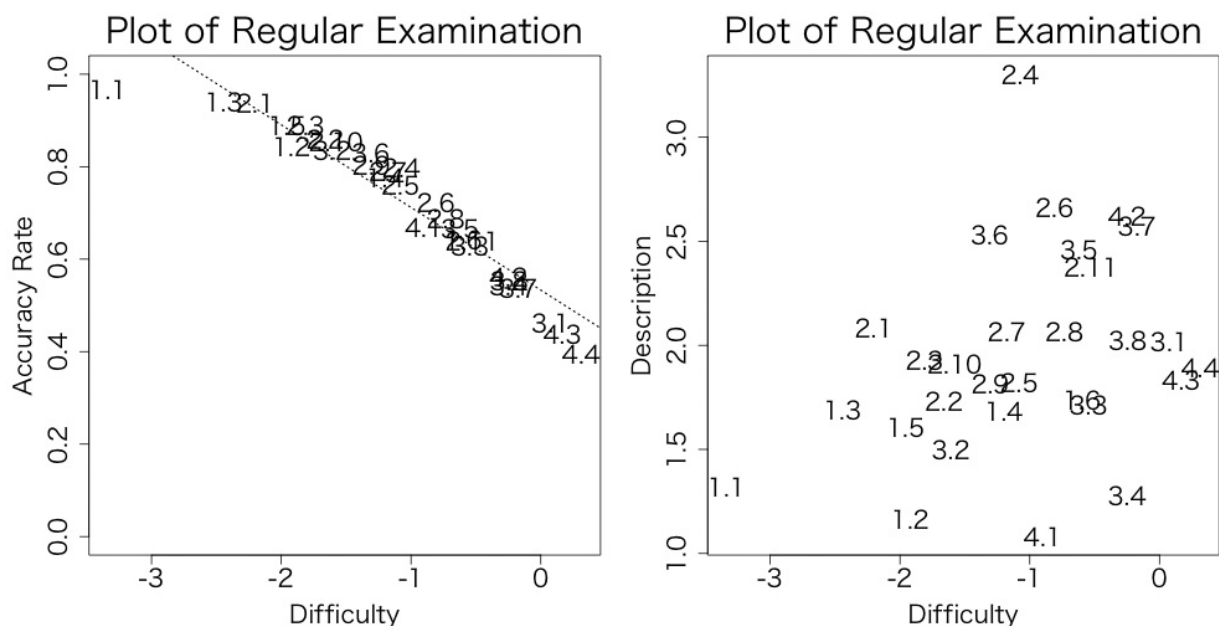


Fig. 3. Scatter Plots of Parameters of Regular Examination under IRT

らの山も正規分布している様子がうかがえる。すなわち学力層が大きく分けて2つあることを意味し、学力の多様性が示唆される。定期試験については29点満点で、平均は21.0、標準偏差は7.18であった。こちらは一見して正規分布とは見えず、むしろある程度の区間に限ってみると一様分布に近いといえる。26点以上についてはクラス分け試験と同様に、上限を定めているために固まって分布したものと思われる。

### 3. 結果に関する考察

#### 3・1 IRT を利用した解釈

クラス分け試験と定期試験のどちらに関しても、Rを利用してIRTを施し、項目のパラメータおよび受験者の特性値を考察した。項目のパラメータは、困難度と識別力である。クラス分け試験に関して、困難度と正答率の、および困難度と識別力の散布図が図2である。困難度と正答率には非常に高い負の相関があり、回帰直線から大きく外れた項目がなく、15問のいずれの項目も別段の問題がないといえる。もし大きく外れる問題があった場合は、出題者が意図しないなんらかの（ひっかけ問題のような）難しさがひそんでいる可能性がある。

困難度と識別力の図においては、識別力が非常に高い問題が見られる。このうち項目08は、以前に他の試験で高い識別力があるという結果が出たのだが、今回の項目にあつては、それほど高く出なかったのが意外であった。他の試験における項目08の識別力は3.5であったが、他の試験では相対的に非常に高く推定されたものの、今回の項目としては他に識別力が高いためあまり目立たないということかもしれない。加えて、今回での識別力も3前後ということで、受験者集団が変わったにもかかわらず、それほど大きな差がないことを、IRTの妥当性と見ることができる向きもある。

同様に定期試験に関する散布図が図3である。正答率と困難度の散布図を見ると、1.1が多少、回帰直線から離れているが、これは正答率が高すぎるためと言えるし、他の問題はほぼ回帰直線付近に分布しており、こちらも問題が妥当であったと考えてよい。興味深いこととして右図において、困難度が各分野ごとにある程度まとまって分布している点がある。各分野（1.1～1.6が文字式、2.1～2.11が指数と対数、3.1～3.8が三角関数、4.1～4.4が応用問題）ごとに、受験者の得手不得手を反映しているようで不思議である。

定期試験の困難度と識別力の散布図からは、困難度が若干低い項目が多かったが、全体的にはやはり妥当な問題であったと判断できる。識別力が最も高い項目2.4の内容は $\sqrt[3]{2}\sqrt{4}$ であり、計算が簡単のため純粋に『定義を正確に把握しているか』を問うことになるので、高い識別力となったものと考えられる。それに対して項目2.5の問題は $\frac{\sqrt{64}}{\sqrt{2}}$ という、分数や2桁の計算を含む複合的な尋ね方をしているため、識別力が項目2.4に比べて相対的に落ちるのであろう。

本稿では2つの試験の項目に加えて、受験者の特性値の比較を試みた。単純に2つの試験の点数に関してプロットしたものが図4の左である。重複する点が重なって表示されるため見づらいが、概ね直線で近似でき、実際相関係数も0.787と十分高い。これを得点ではなく、IRTで推定された特性値に関してプロットしたものが図4の右である。こちらは

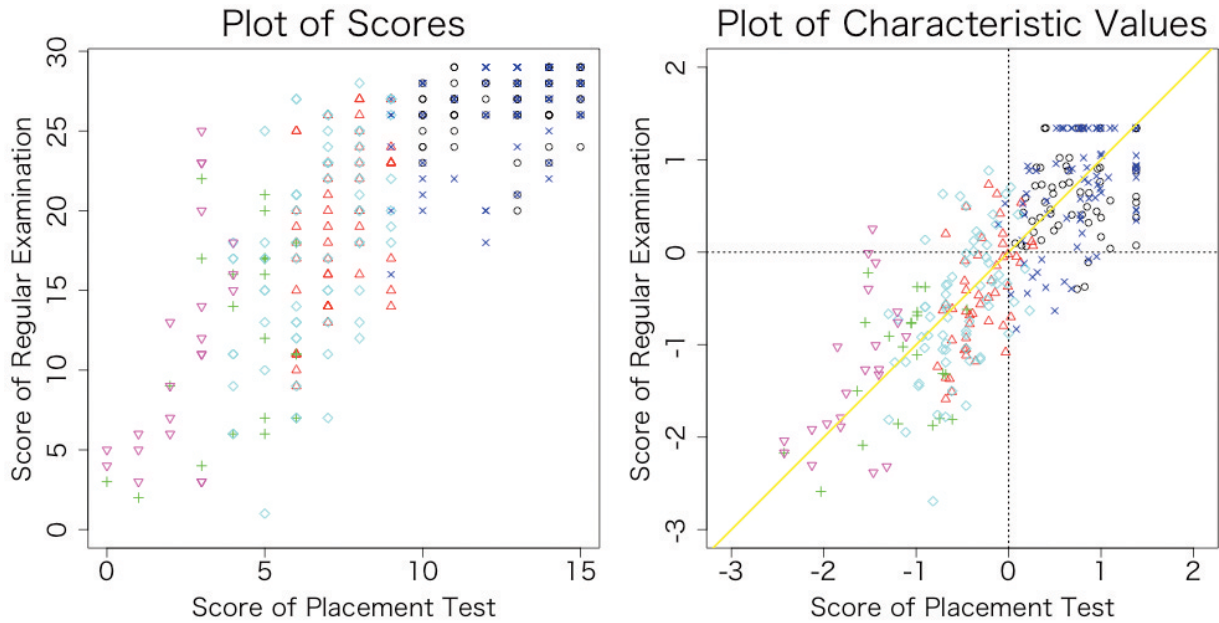


Fig. 4. Scatter Plots with respect to Scores and Characteristic Values of our two Tests

実数値でプロットするため、完全な重複は少なく、より見やすい。どちらの図も習熟度別のクラス分けによって色づけをしており、本稿には掲載しないが各クラスごとの散布図を作成した。それらの特徴からクラス分けの効果を検討してみると、6クラスのいずれの図においても、

- (1)  $y = x$ の右下（特性値が落ちた場所）にいる点が極端に多いわけではない、
- (2) どのクラスでも、 $y = x$ の左上（特性値が向上した場所）に点が存在する、
- (3) どのクラスでもシャピロ・ウィルク検定によって、ほぼ正規分布だと解釈できる、

という状況が見受けられる。すなわち受験者に適切な教育が提供できた根拠としたい。仮にクラス分けをしていない場合、もっと右下の点が増えたり、左上に点が増えない、という結果になりうる。

### 3・2 講義担当者の意見

本節では、データでは見えないクラス分けの効果を把握するために、クラス分け後の講義担当者の意見を記述する。前年度までの、クラス分けをしない講義も担当経験のある者からは、層分けができて講義の照準を定めやすくなるため



Fig. 5. Scatter Plots of Scores at similar two classes



よい、との意見がある。つまり(1)講義内容が既知である層、(2)講義内容が未知だが学習意欲のある層、(3)講義内容が未知で学習意欲のない層、という存在を感じるが、クラス分けをしなければ一斉にその対応をしなければならず、どこに合わせても（少なくとも担当者にとっては確実に、あるいは受講者にとっても）消化不良の感が残る。クラス分けをしないことによる人数の多さもあいまって、多様な学生への対応は困難である。また今期からの担当者の意見も概ね同様で、習熟度別に分かれているために、講義内容の目処をたてやすく、授業と演習の配分がスムーズであった、などが挙げられる。

それでも上位の学生たちには多少退屈な講義があったかもしれないとの反省の声もあり、学習意欲の高い学生の退学防止という観点でも、もっと細かいクラス分けを実施することが望ましい。実際、基礎クラスこそ30名弱の少人数制といえる程度の人数であったが、標準クラスおよび発展クラスは50名～80名ほどであり、密な接触が期待できるとは言い難い。なお演習の必要性に関しては、複数の著者で概念の把握（覚えること）と理解（意味を知ること）の差を埋めるために不可欠であるとの共通認識をもっている。層分けにより与えるべき課題がより明確になることは必然であり、演習時間の確保のためにもクラス分けは有効である。

また標準クラスに関しては2クラスあったが、同一の担当者で人数のみが異なる（49名と78名）状況であった。図5はそれらにおける2つの試験結果の散布図である（2つの試験のいずれか一方の結果が欠けているものは除外したため、図の点は48個および68個となっている）。クラス分けテストの点数の差はほほない状況であったのに対し、定期試験の点数に関しては、49名のクラスの方がばらつきが少なく（49名のクラスでは $SD=5.00$ 、78名のクラスでは $SD=6.44$ ）、図からも人数の多いクラスよりも少ないクラスの方が適切な講義ができたかと主張できる。

#### 4. まとめ

習熟度別のクラス分けに関しては一般的に、クラス分けの手間や人的コストを要するが実施する意義は大きい、として差し支えないであろう。本稿では実際に、試験の結果と担当者の声をもって、その考えを実証できた。特に図5のように、よく似た学力層で人数のみ多寡のあるクラスを同一の担当者が担当する、という状況をはからずも設定できたが、これは人数による影響がはっきりと差が出た好例だと認識している。数学を必要とする工学において、学生がその基本的な内容を修めるために、クラス分けはやはり必要であると結論づけたい。

#### 文 献

- (1) 中嶋康博, 佐々木良勝, “項目反応理論による数学基礎力確認テストの解析”, 久留米工大研究報告, No. 39 (2016), pp. 63-71.
- (2) 豊田秀樹, “項目反応理論 [入門編]” (2012), 朝倉書店.
- (3) 加藤健太郎, 山田剛史, 川端一光, “Rによる項目反応理論” (2014), オーム社.
- (4) 植野真臣, 莊島宏二郎, “学習評価の新潮流” (2010), 朝倉書店.
- (5) Dimitris Rizopoulos, “ltm: An Package for Latent Variable Modeling and Item Response Theory Analyses”, Journal of Statistical Software, November 2006, Volume 17, Issue 5.
- (6) Dylan Molenaar, Francis Tuerlinckx, Han L.J. van der Maas, “Fitting Diffusion Item Response Theory Models for Responses and Response Times Using the R Package diffIRT”, Journal of Statistical Software, August 2015, Volume 66, Issue 4.