〔論　文〕

項目反応理論による数学基礎力確認テストの解析

中嶋　康博[*1]・佐々木良勝[*1]

Analysis of Basic Mathematics Examination under Item Response Theory

Yasuhiro NAKASHIMA[*1], Yoshikatsu SASAKI[*1]

Abstract

Many students may have difficulty studying mathematics during their first year at our university. Mathematics is an essential course, and we intend that the difficulty should be reduced suitably for each characteristic. To achieve this, we needed to investigate the students' basic mathematical knowledge and calculating ability. Therefore, we administered an examination. Their scores were informative. Moreover, we used Item Response Theory by R-language in assessing the test. The theory aids us in checking whether certain examination items were appropriate. This report contains figures that represent the characteristics of the items of our examination, as well as our assessment of the results.

Keywords：Item Response Theory, Mathematical Education, First Year Education, Remedial Education, R-language

1．緒　　言

大学全入という言葉を耳にして久しく，本学でも複数の入学試験のもと多様な学生が在籍しており，特に高校時代に履修した学習内容に関して，入学時点で差異がある学生が混在している．高校の工業科出身の学生であれば数学Ｉまでの履修ということもあるが，そのような学生は工業系の技術を修得していることも多く，普通科出身の学生は実習面で遅れをとる．この点では学生同士で得手不得手な内容を相互に教え合うなど，補完の効果も期待できる．しかし数学の初年次教育を担当するものとしての所感を述べると，学生の基礎的な数学の能力が多様なため，一様な授業展開が学生の学習意欲や理解を阻害する可能性を感じる．そのため各人に応じたクラス編成をできると望ましい．学生によっては，意欲はあるもののこれまでに学習機会が少なかったため，基本的な定義が定着していないというケースも，実際にある．工学を専攻する以上，学生に数学や物理は必須であり，彼らに適切な機会や内容の提供が必要である．

このような背景の中で今回，初年次の学生を対象とした，数学に関する基礎的な能力の調査（数学基礎力確認テスト）を行った．これにより受験者の得点や正答率などの基本的な情報を得たが，著者はテスト理論に興味があり，この結果に項目反応理論の適用を試みた．『テストを作成・実施・評価・運用するための実践的な数理モデル』である[(1)]とされる項目反応理論による結果の解釈が，本学の現状理解とよりよい学習支援につながることを期待する．本稿は次節で項目反応理論に関して説明し，その後，数学基礎力確認テストの説明および項目反応理論による解釈を行う．実施したテストおよびその各問の結果は付録に記載した．その結果得られた知見としては，テストの問題が十分妥当な内容であったこと，識別力が高くなる問題の傾向を推測できたこと，学生の多様性が確認されたこと，などが挙げられる．

2．項目反応理論

本節では，数学基礎力確認テスト（以下，試験と表記）の解析に利用する，項目反応理論（IRT，Item Response Theory，項目応答理論とも呼ばれる）について言及する．用語等の多くは豊田[(1)]に準拠する．本稿で頻出する用語に関しては，試験を受験した学生を受験者，試験における合計点を得点，推定された受験者のパラメータを特性値，試験の各問題を項目，などと表すこととする．

[*1] 教育創造工学科
平成28年12月７日受理

IRT は『テストの尺度化とテストの等化』[2]を目的とするテスト理論の中で発展した手法で，古典的テスト理論（CTT, Classical Test Theory）と対比される．欧米ではスタンダードな理論として各種試験に利用され，日本では TOEIC や基本情報技術者試験などに利用されている．その理由として，次のような利点があげられる[1]．

・複数のテスト間の結果の比較が容易であること．
・測定精度をきめ細かく確認できること．
・平均点をテスト実施前に制御できること．
・テスト得点の対応表が作成できること．
・受験者ごとに最適な問題を瞬時に選び，その場で出題できること．

実際にこれらの恩恵を受けるには，等化という手法の利用や計算機を利用した出題などが必要となるが，本稿ではそれらは行わず，主として項目評価のために用いた．個人的には上記に加えて，後述する項目特性曲線（ICC, Item Characteristic Curve）などによる，視覚的に理解しやすい点も IRT を利用するメリットであると感じる．

IRT においては，項目をロジスティック曲線で表す．通常そのパラメータは 1 個から 3 個とされるが，植野・荘島[3]によれば 4 個や 5 個のパラメータをもつモデルが示されている．ここでは本稿で利用した，２つのパラメータをもつモデル（2PLM, 2-Parametric Logistic Model）を説明する．2PLM におけるロジスティック曲線は，受験者の特性値を θ，その項目に正答する確率を $P(\theta)$ として，２つのパラメータ α と β を用いて

$$P(\theta) := \frac{1}{1+\exp(-D\alpha(\theta-\beta))}$$

と表される．これは標準正規分布の分布関数

$$F(x) := \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

によく似た曲線であり，積分計算の都合で正規分布よりも便利な関数として扱われる．ロジスティック曲線の式自体は『表現したいもの，すなわち項目の特性』に合わせて作られているため，きれいな式であるとは表現しづらいが，この式で表される曲線の特徴がよい．式において D は，標準正規分布の分布関数との近似に必要な定数であり，その近似の精度のために通常 $D=1.7$ が利用される[1]ため，本稿でも $D=1.7$ としている．α は識別力と呼ばれ，グラフでは ICC の立ち上がりの緩急を表し，項目としては受験者の特性値を判定しやすい指標，すなわち項目の『よさ』を表す．通常 $0.3 \leq \alpha \leq 2$ 程度で推定されるといわれる．β は困難度と呼ばれ，グラフでは ICC の横軸方向のスライドを表し，項目としては難易度の指標，すなわち高いほど項目が『難しい』ことを表す．通常 $-2 \leq \beta \leq 2$ 程度で推定されるといわれる．

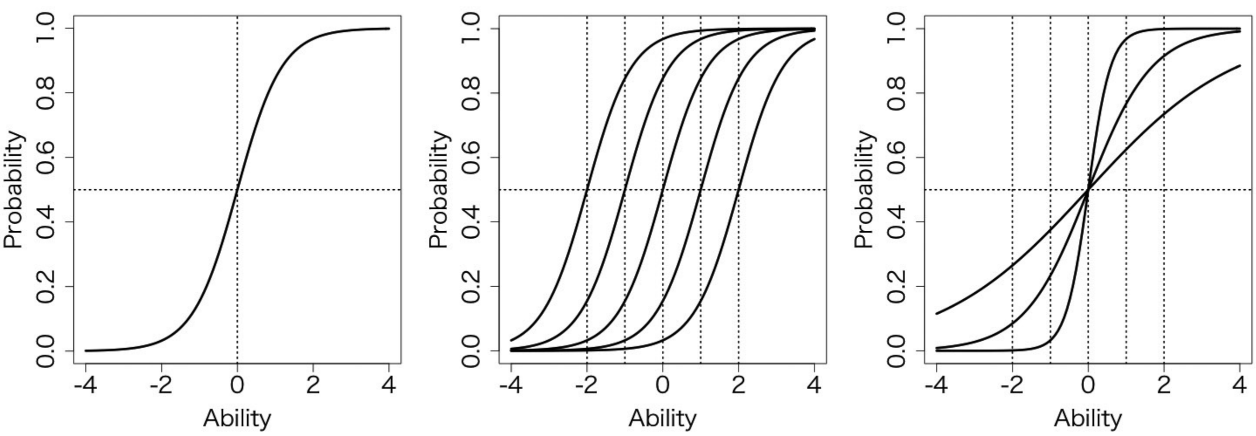


Fig. 1. Examples of Item Characteristic Curve

図 1 は，パラメータを変更して 2PLM のロジスティック曲線を描いたものである．各項目ごとに 1 つの曲線が描かれ，横軸に受験者の特性値 θ をとり，縦軸にはその項目の正答率 $P(\theta)$ をとる．左図がもっとも標準的な $\alpha=1$，$\beta=0$ の曲線であり，特性値 θ が 0 である点に関して対称になる．またちょうどそのとき正答率が0.5である．ロジスティック曲線は標準正規分布の近似であることによれば，たとえば $-1 \leq \theta \leq 1$ の範囲には68.3%のデータが存在し，$-2 \leq \theta \leq 2$ の範囲には95.4%のデータが存在することになる．偏差値の考え方でいえば，たとえば特性値が $\theta=1$ である状態

が（受験者集団における）偏差値60に対応する．中央の図は，$D=1.7$，$\alpha=1$として，βを-2から2まで1ずつ変化させたものである．5本の曲線があるが，左から順にβが-2，-1，0，1，2の曲線である．困難度とは，ちょうどその特性値において正答確率が0.5になるような曲線であり，識別力αが1でない場合でも同じである．また右図は$D=1.7$，$\beta=0$として，αを0.3，0.7，2と（特に意味はないが描写の都合で）変化させたものである．もっともなだらかな（$\theta=-4$でも0.1以上の正答率がありそうな）曲線が$\alpha=0.3$であり，ちょうど3本の間に位置するものが$\alpha=0.7$，もっとも急な（$\theta=1$でほぼ100%の正答率がありそうな）ものが$\alpha=2$である．αの値が高ければ，ある一定の特性値θ_0（これが困難度βになるが）までは正答率は低いが，特性値がθ_0の前後で急激に正答率が上がる．したがって受験者の特性値を判断するのに都合がよい指標であるといえる．

以上がICCに関する説明であるが，IRTの適用に関する1つの仮定について述べる．一般にテストにおける項目では，複合的な知識が必要となることがある．たとえば『2進法で表示された数を10進数で表示せよ』といった項目の場合，これは数学の問題であるが情報の知識を含んでいるともいえる．また複雑な文章題であれば，問題文を読み取る力が暗に要求されるかもしれない．これらのような項目を複数含むテストで因子分析を行うと，2つ以上の因子があると判断される可能性がある．つまりテストで測定しようとする能力に2つ以上の傾向があることになるが，IRTにおいてはそのモデル作成の考え方から，特性値θのみによって正答率が定まる必要があるため，因子が2つ以上あると考えられる状態は望ましくないとされる．言い換えればテストで測っている因子が1つだけである必要性があり，測定の一次元性といわれる[2]．一次元性の確認には，因子分析の実施やスクリープロットの作成が有効であるが，本稿ではスクリープロットを作成したところ，図2のようになった．図から判断して，第1因子の固有値が極めて高く，第2因子以降はいずれも第1因子と比べると十分に低い．このため我々の試験は一次元性をみたしていると判断でき，項目の選定が妥当であったといえる．なおどのような項目でテストの一次元性がくずれるのかという点にも興味があり，今後たとえば上述のような項目を意図的に配置し解析するような課題がある．

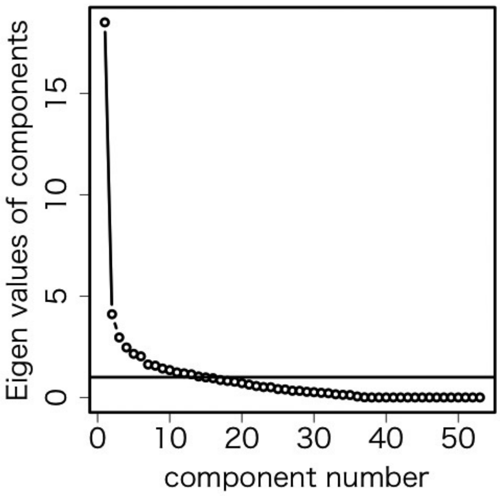


Fig. 2. Scree Plot of the result of our examination

3．数学基礎力確認テスト

3・1　テストの概要と結果

本節では実際に著者が使用した試験の概略とその結果，およびIRTによる解析を記述する．用語や略語などについては前節のとおりとする．試験は2016年度の1年次に開講される講義にて，受講者の現状を把握する目的で実施した．全5学科中，1学科Aの半分と残り4学科全体の，293名が受験した．ただしAと異なる1学科Bを除いた4学科では，一度に実施することによる学生の負荷を考慮し，4日間の4回に分けて20分程度ずつ実施した．このため欠損値が生じ，結果や解析に使用した有効なデータは226名分となる．試験は53問からなり，いずれも基礎的な数学の内容である．工学部に所属する学生が受験者である点を考慮して，図形の問題を比較的多く取り入れた．具体的な問題は付録に記載した．各項目はいずれも記述式で，部分点はない．つまり公式や解法は理解していても計算間違いをした場合は得点にならず，そのため今回の試験では計算力の高い学生が高く評価される傾向にあると考えられる．すべての項目において項目間の依存はなく，ある項目での誤答が他の項目の誤答を誘導することはない．得点に関しては重みをつけず，

正答は１，誤答は０として53点満点で評価した．各項目の難易度には当然ながら差があると考えられるが，その評価(得点の重み）に IRT を用いるのであり，また本稿の解析で利用した R の ltm パッケージではその差をあらかじめ評価に反映させる必要はないためである．

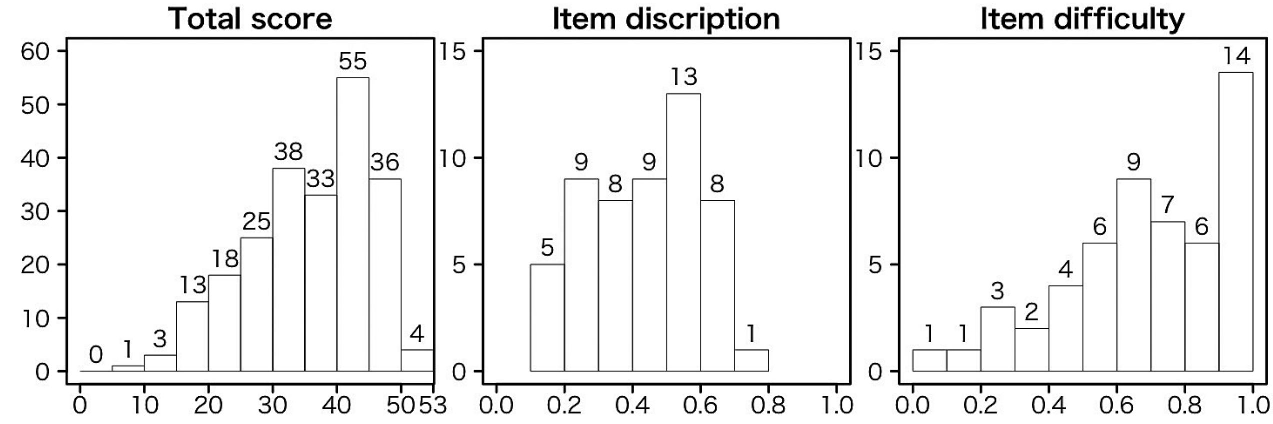


Fig. 3. Histograms of total score, item description, and item difficulty under CTT, of our examination

試験の結果は，得点の平均が36. 3，標準偏差が9. 5であった．最高点は51点，最低点は７点である．各項目の正答率も基本的な情報にあたるが，正答率は古典的テスト理論（CTT）では困難度と表現されるため，補足として CTT によるデータも追記する．図３は左から順に試験の得点，CTT における識別力，および CTT における困難度の分布をそれぞれ表す．得点に関する図によると，53点満点であるため右側の裾の尾が切れているが，40点～45点の層を中心に正規分布にしたがうようにも見える．また平均値と中央値はほぼ同じくらいであり，最頻値がそれらよりも若干高めに現れているものの，やはり正規分布に近い分布といえそうである．受験者にとって，比較的容易な項目が多かったといえるかもしれない．しかし満点のおよそ半分である25点を割り込む層が，人数にして35名程度の15％ほど存在しており，学力の多様性がうかがえる．図３の中央の図に関して，CTT での識別力は各受験者の正誤と総得点の相関係数として定義される．加藤・山田・川端[2]によればこれは，『ある項目への正誤がテスト得点と高い正の相関を持てば，その項目へ正答した受験者はテスト得点自体も高得点である傾向』があり，IRT における識別力と同じ意味合いの指標である．識別力は，図よりほぼ一様に分布しているといえるが，IRT における識別力ではその分布を見ることで，どの層に適したテストであったかを判断できるのと比べて，CTT では識別力が一様な分布であったという以上の知見が得られるかどうか，不明である．図３の右図について，CTT における困難度については項目の正答率のことで，高いほど容易な項目であるといえる．困難度が0. 9，つまり正答率が90％を超える項目が14もあったが，基礎的な内容の確認を目的とする本稿では妥当な結果といえる．当然のことであるが，ヒストグラムに基づき平均値を計算すると36. 2となり，ほぼ実データの平均得点となる．CTT による困難度と識別力の２つの指標に関してデータを見ると，たとえば項目１では困難度が0. 987で識別力が0. 15であった．一般的な解釈として，極端な困難度をとる（すなわち容易あるいは難解な)問題では識別力は落ちるが，項目１はその解釈にしたがう結果である．逆の例では項目33が，困難度は0. 035でもっとも難しく，しかし識別力が0. 237であった．難しい問題であれば，それが解ける時点で総得点は高いであろうから，識別力はそれほど落ちない，ということかもしれない．

３・２　IRT の適用結果

本稿ではモデルに 2PLM（2Parameter Logistic Model）を採用した．選択式の解答がなくすべて記述式であったためあて推量を伴う 3PLM は不要だと思われるし，2PLM であっても計算機による処理に特段の問題を感じなかったので，計算上の利点しかないと思われる 1PLM を利用する必要がなかったためである．処理は統計処理用のソフト『R』を用い，項目反応理論向けのパッケージである ltm パッケージを利用した．パッケージ名が IRT ではないが，ltm とは Latent Trait Model のことであり，豊田[1]によれば，Latent Trait Theory（潜在特性理論）を項目反応理論と呼び直し，こちらの方が定着したとされる．本稿では226×53の２値行列を反応パターン行列とし，ltm パッケージの関数である ltm 関数を用い，能力の周辺分布を利用した最尤推定法により各項目の２つのパラメータを推定した[2][6]．なお R の ltm クラスではその要素 coefficients に２つの特性値が書かれているが，そのうち intercept の列は本質的には困難度のことであるが，困難度そのものを表してはいない点に注意が必要である．ltm クラスに factor.scores 関数を施して得られる

fscores クラスの要素 coef には識別力と困難度が直接記述されている．

各項目に対して推定された 2 つのパラメータによって，ICC と散布図を描いたものが図 4 である．項目が多いため ICC が重複し見づらい面があるが，ltm クラスに対し plot 関数を施すと ICC が表示されるのは便利である．もしも一部の項目だけで ICC を描きたいのであれば，ltm クラスの coefficients を部分的に書き換えればよい．色付けに際しては便宜的に項目を分類しており，分数，２次元図形，３次元図形，文字式，文章題，その他に対してそれぞれ，赤，橙，緑，水色，藍色，桃色とした．

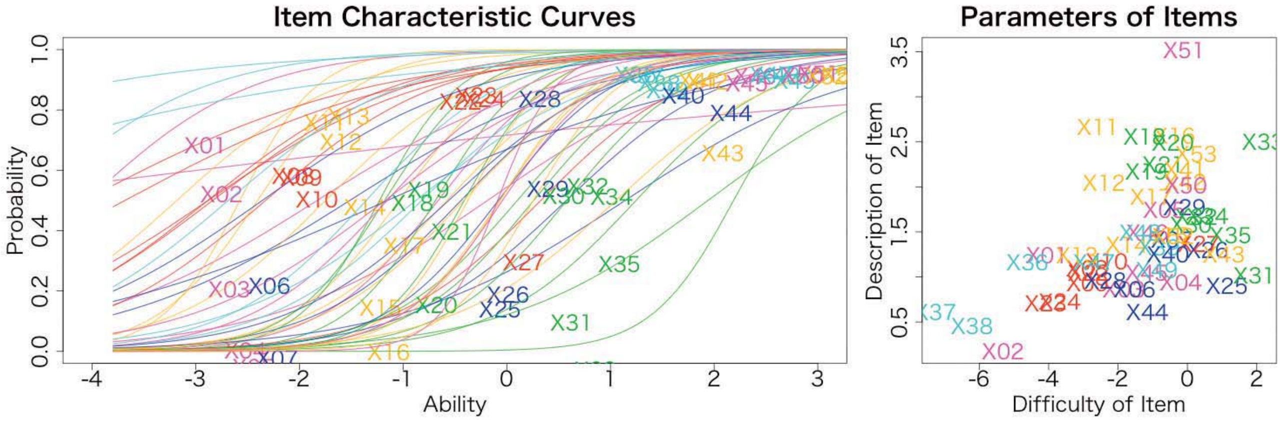


Fig. 4. Item Characteristic Curves, and Scatter plot parameters of our examination

IRT によるテストの解釈について言及する．識別力と困難度については項目ごとの数値を付録に記載した．各項目の特徴を把握するために，IRT における識別力と困難度の分布を，図 3 の得点の分布を再掲して図 5 とする．識別力が2.5を超えるのは11，16，18，51の 4 項目があり，低いのは項目 2 が0.174だったが，それ以外はすべて0.3を超えていた．困難度が 1 を超えるのは25，31，33，35，43の 5 項目であり，逆に（全体的に低かったがその中でも特に）低いものでは 2，23，36，37，38などが挙げられる．IRT を利用する際には（そして一般的なテストにおいてもだと考えられるが）極端な正答率をとる項目は解析にはふさわしくない．困難度と正答率を考慮すると，たとえば項目31や33は難しすぎるし，項目 2，23，36，37，38などは簡単すぎる上に識別力が低いため，これらは試験から除外する案が考えられる．

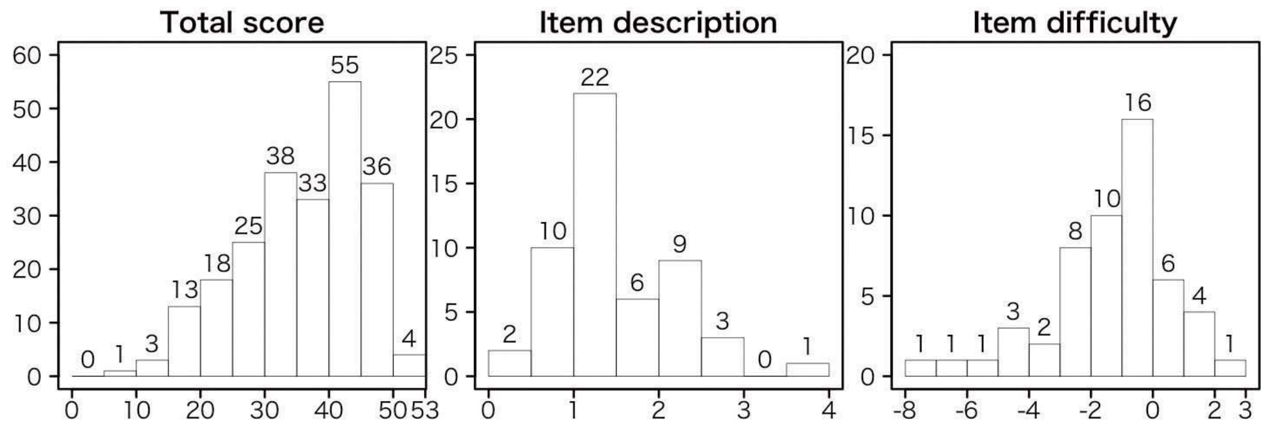


Fig. 5. Histograms of total score, item description, and item difficulty under IRT, of our examination

CTT と IRT の比較として図 6 を作成した．左と中央がそれぞれ CTT と IRT のパラメータの散布図であり，右は困難度に関して横軸に CTT を，縦軸に IRT をとって散布図を描いたものである．左図においては，CTT の困難度すなわち正答率を－1 倍して反転させて，横軸が増えるほど難しいように変換している．左と中央のどちらの散布図においても，項目51は困難度がほどほどで識別力が高い．また項目33は，困難度はどちらにおいても最も高い一方で，識別力に関しては随分差があるように見える．色合いや位置関係などを見ると，おおまかな傾向は両者で共通している気もするが，たとえば（3 次元図形と分類した）緑の項目群は CTT よりも IRT の方で総じて難しいと判断されている．

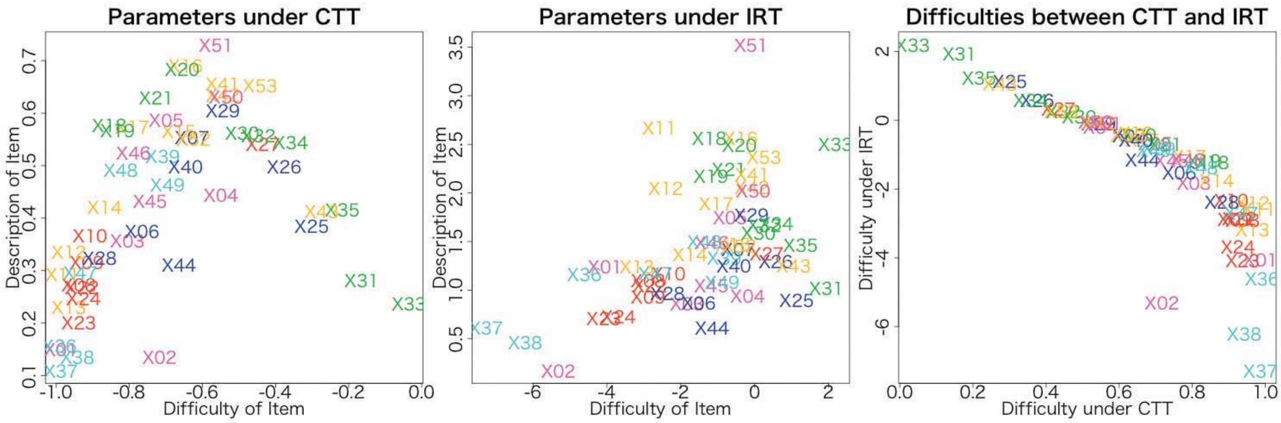


Fig. 6. Comparison between parameters under CTT and IRT of our examination

右図においては横軸つまり正答率が上がるほど IRT での困難度が低くなる傾向があることを示す図であるが，項目２がその傾向からはずれている．質問内容が（一見すると）よく似た項目１と２において，正答率では項目１の方が高いが，困難度で見ると項目２の方が簡単だと判断されている点が興味深い．項目２に関しては，正答率は72%，困難度は－5.31，識別力は0.17であった．つまり項目２は一部の人は間違っているのだが，非常に容易な問題だと判断され，特性値を見るにはほとんど機能しない，と解釈できる．実は項目２はある種のひっかけ問題になっており，著者も出題時にそのように認識していた．出題がふさわしいかどうかはともかく，出題時の認識と相違ない結果となっており，IRT の妥当性を感じる．６図の右図において，項目２以外は極端に傾向が異なる項目は存在せず，２節で見た一次元性と合わせて出題した項目に不備が少なかったことが確認できたと考える．細かい点では，項目36と37において，質問自体はほぼ同じ内容であり正答率も同じ99%であるのに，困難度が異なり，項目37の方が容易だと判断された．項目37は，解釈によってはひっかけ問題とも取れるが，この点で項目２と似た結果になったと考える．

項目31と33は，どちらも図形の表面積を問う問題で，正答率，困難度のどちらで見ても，難しい問題の１，２番目だった．しかし識別力で見ると，随分と差があり，この解釈は難しいが，項目33はより計算力が問われる内容であったためかと考える．すなわち，本質的には同じ内容を問う２問において，より計算力が必要となる問題ほど識別力も高くなるのではないか，と推測した．

著者の意図に反した結果としては，項目51が挙げられる．項目33のような難しすぎる問題では正答率が下がり識別力が落ちるのはわかるが，項目51のような普通の問題（正答率56%）において，最大の識別力が出るのが意外であった．項目51の困難度はほぼ０であり，たとえば項目29は困難度がほぼ０で正答率は54%とよく似ているが，識別力は1.78である．項目51については，定義がそのまま問題になったようなものであり，このような項目はブレが少ないであろうから当然高い識別力をとる，ということであれば，IRT は問題の内容は考慮していないにも関わらず識別力のイメージに適する結果を返している点で，IRT の有効性を示す結果といえる．たとえば出題者が定義に近い内容を問うつもりで出題したが，IRT の結果であまり識別力が高くないような状況が起きた場合に，その問題を見直すといったテストの設計を振り返るのに利用できるし，そもそもそれが IRT の本来の利用方法であろう．IRT から得られる知見によれば，また内容の面での判断が不可欠であるという前提はつくが，受験者とのやり取りを含めて項目の比較をすることで，受験者の現状に適した教育を考察できるかもしれない．若干の留意点としては，正答率で見れば95%以上の問題が複数あるが，それらは困難度に関して比較的バラけて配置されている点である．たとえば項目１と項目11において，正答率の点では99%と98%でほぼ同じ程度であるが，困難度は順に－4.05と－2.57であり，項目11のほうがより難しいと判断されている．

４．ま と め

今回の処理を経て IRT による処理結果の妥当性を度々実感することができ，また出題した項目に関して適切かどうかを考察する材料を得た．識別力に関しては，それを高めるための推測ができたため，今後あらためてその推測を考察したい．識別力が高い項目を幅広い困難度の項目の中から集めることにより，その試験で受験者の特性値を少ない誤差で測定できるという点で，識別力の高い項目の作成は有意義なためである．実際の解析においては，R とそのパッケージを利用すると非常に容易に項目のパラメータが得られ，R と IRT の高い利便性が確認できた．なお本稿では考察に

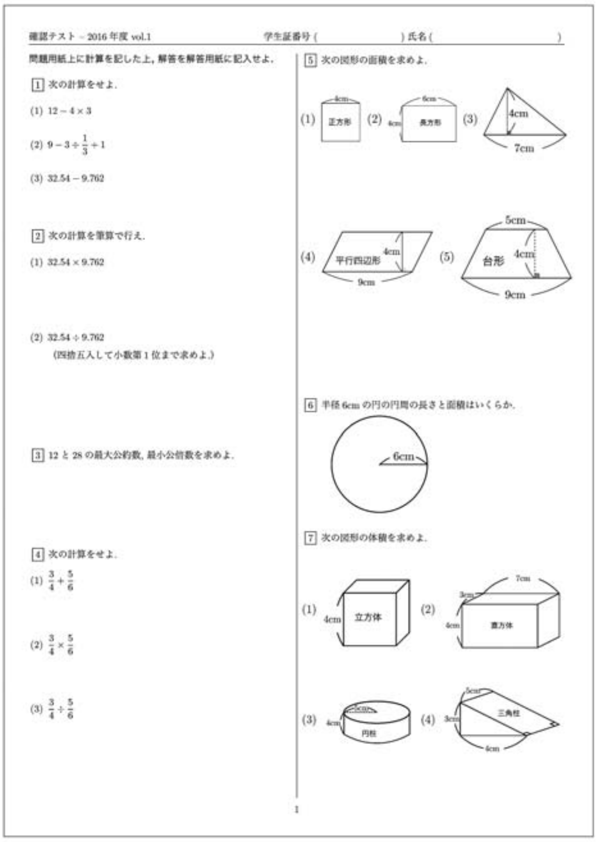
至らなかったが，取り組むべき課題も多く見つかった．たとえば本学では物理教育において先行の研究がなされており[6]，本稿の内容と合わせて知見が得られるかもしれない．またIRTで得られた結果に対する連関規則や二分木探索などのデータマイニングも効果を期待している．多様な学生により適した学習支援を提供すべく，これらの経験を活用したい．

5．謝　辞

本稿の作成にあたり，テスト理論や項目反応理論に関する貴重な意見をいただいた，久留米工業大学共通教育科堀憲一郎先生に感謝いたします．

5．付　録

5－1．数学基礎力確認テスト



確認テスト － 2016年度 vol.1　　学生証番号（　　　　）氏名（　　　　）

問題用紙上に計算を記した上，解答を解答用紙に記入せよ．

1 次の計算をせよ．

(1) $12-4\times 3$

(2) $9-3\div\frac{1}{3}+1$

(3) $32.54-9.762$

2 次の計算を筆算で行え．

(1) 32.54×9.762

(2) $32.54\div 9.762$
（四捨五入して小数第1位まで求めよ．）

3 12と28の最大公約数，最小公倍数を求めよ．

4 次の計算をせよ．

(1) $\frac{3}{4}+\frac{5}{6}$

(2) $\frac{3}{4}\times\frac{5}{6}$

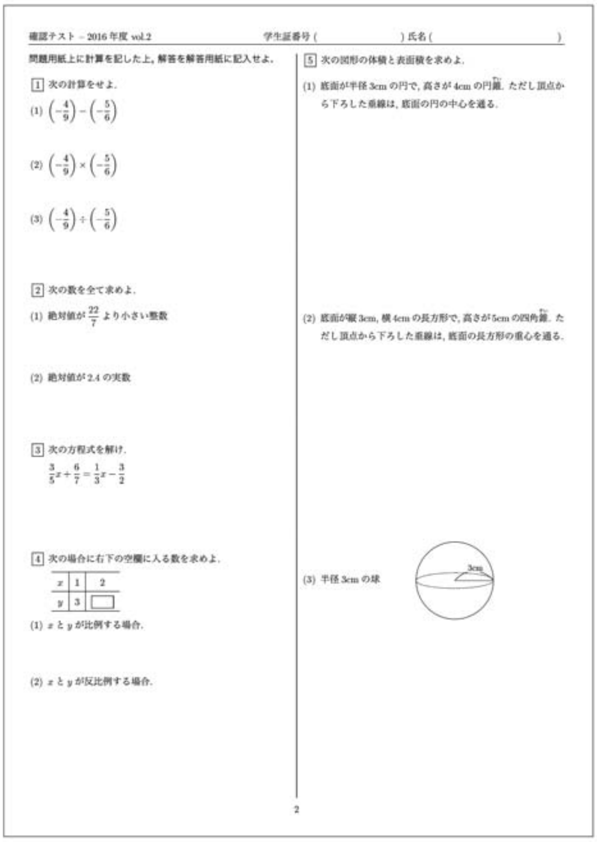
(3) $\frac{3}{4}\div\frac{5}{6}$

5 次の図形の面積を求めよ．

6 半径6cmの円の円周の長さと面積はいくらか．

7 次の図形の体積を求めよ．

1



確認テスト － 2016年度 vol.2　　学生証番号（　　　　）氏名（　　　　）

問題用紙上に計算を記した上，解答を解答用紙に記入せよ．

1 次の計算をせよ．

(1) $\left(-\frac{4}{9}\right)-\left(-\frac{5}{6}\right)$

(2) $\left(-\frac{4}{9}\right)\times\left(-\frac{5}{6}\right)$

(3) $\left(-\frac{4}{9}\right)\div\left(-\frac{5}{6}\right)$

2 次の数を全て求めよ．

(1) 絶対値が $\frac{22}{7}$ より小さい整数

(2) 絶対値が2.4の実数

3 次の方程式を解け．

$\frac{3}{5}x+\frac{6}{7}=\frac{1}{3}x-\frac{3}{2}$

4 次の場合に右下の空欄に入る数を求めよ．

x	1	2
y	3	□

(1) x と y が比例する場合．

(2) x と y が反比例する場合．

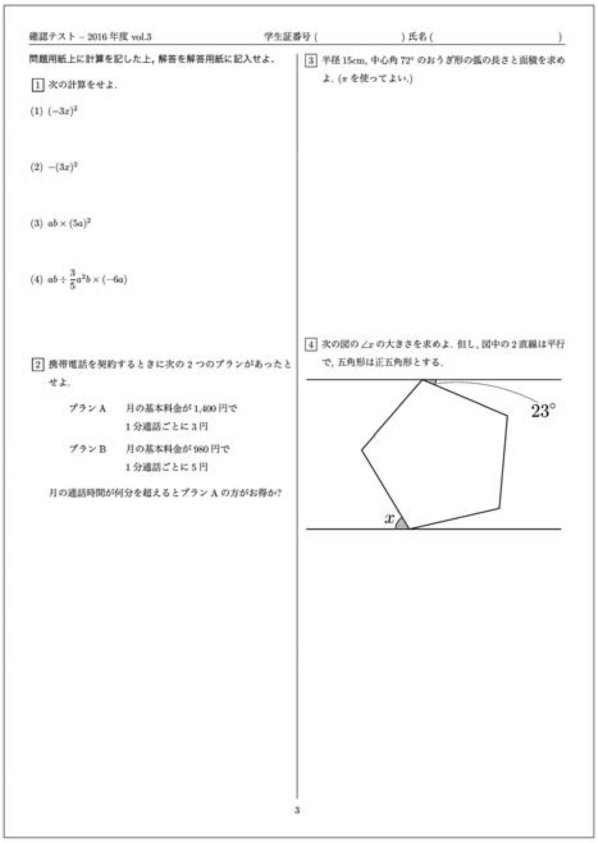
5 次の図形の体積と表面積を求めよ．

(1) 底面が半径3cmの円で，高さが4cmの円錐．ただし頂点から下ろした垂線は，底面の円の中心を通る．

(2) 底面が縦3cm，横4cmの長方形で，高さが5cmの四角錐．ただし頂点から下ろした垂線は，底面の長方形の重心を通る．

(3) 半径3cmの球

2



確認テスト – 2016 年度 vol.3　　学生証番号 (　　　　)氏名 (　　　　)

問題用紙上に計算を記した上，解答を解答用紙に記入せよ．

1 次の計算をせよ．

(1) $(-3x)^2$

(2) $-(3x)^2$

(3) $ab \times (5a)^2$

(4) $ab \div \frac{3}{5}a^2b \times (-6a)$

2 携帯電話を契約するときに次の 2 つのプランがあったとせよ．

プラン A　月の基本料金が 1,400 円で 1 分通話ごとに 3 円

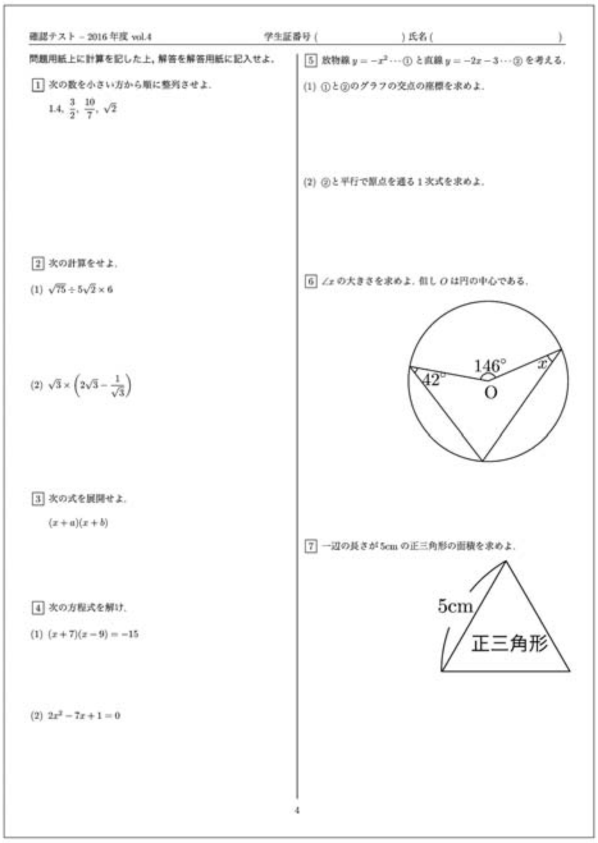
プラン B　月の基本料金が 980 円で 1 分通話ごとに 5 円

月の通話時間が何分を超えるとプラン A の方がお得か?

3 半径 15cm，中心角 72° のおうぎ形の弧の長さと面積を求めよ．(π を使ってよい.)

4 次の図の $\angle x$ の大きさを求めよ．但し，図中の 2 直線は平行で，五角形は正五角形とする．

3



確認テスト – 2016 年度 vol.4　　学生証番号 (　　　　)氏名 (　　　　)

問題用紙上に計算を記した上，解答を解答用紙に記入せよ．

1 次の数を小さい方から順に整列させよ．

$1.4,\ \frac{3}{2},\ \frac{10}{7},\ \sqrt{2}$

2 次の計算をせよ．

(1) $\sqrt{75} \div 5\sqrt{2} \times 6$

(2) $\sqrt{3} \times \left(2\sqrt{3} - \frac{1}{\sqrt{3}}\right)$

3 次の式を展開せよ．

$(x+a)(x+b)$

4 次の方程式を解け．

(1) $(x+7)(x-9) = -15$

(2) $2x^2 - 7x + 1 = 0$

5 放物線 $y = -x^2 \cdots$ ① と直線 $y = -2x - 3 \cdots$ ② を考える．

(1) ①と②のグラフの交点の座標を求めよ．

(2) ②と平行で原点を通る 1 次式を求めよ．

6 $\angle x$ の大きさを求めよ．但し O は円の中心である．

7 一辺の長さが 5cm の正三角形の面積を求めよ．

4

5－2．IRT によりえられた，各項目の識別力と困難度

項目	困難度	識別力
X1	-4.049	1.243
X2	-5.311	0.174
X3	-1.828	0.865
X4	-0.186	0.945
X5	-0.661	1.752
X6	-1.514	0.868
X7	-0.441	1.421
X8	-2.884	1.099
X9	-2.881	0.933
X10	-2.327	1.175
X11	-2.574	2.675
X12	-2.404	2.054
X13	-3.182	1.241
X14	-1.747	1.366
X15	-0.577	1.472
X16	-0.375	2.569
X17	-1.041	1.895
X18	-1.227	2.565
X19	-1.178	2.176
X20	-0.41	2.491

項目	困難度	識別力
X21	-0.688	2.251
X22	-2.837	1.048
X23	-4.079	0.712
X24	-3.682	0.732
X25	1.142	0.902
X26	0.575	1.301
X27	0.328	1.377
X28	-2.358	0.966
X29	-0.081	1.78
X30	0.102	1.588
X31	1.941	1.019
X32	0.259	1.67
X33	2.188	2.5
X34	0.579	1.684
X35	1.247	1.465
X36	-4.607	1.166
X37	-7.269	0.615
X38	-6.208	0.462
X39	-0.82	1.334
X40	-0.56	1.256

項目	困難度	識別力
X41	-0.064	2.196
X42	-0.069	2.061
X43	1.054	1.257
X44	-1.15	0.612
X45	-1.173	1.051
X46	-1.154	1.495
X47	-2.686	1.172
X48	-1.356	1.497
X49	-0.876	1.082
X50	-0.041	2.022
X51	-0.089	3.523
X52	-0.416	1.456
X53	0.256	2.371

文　献

(1) 豊田秀樹, "項目反応理論［入門編］", 朝倉書店.
(2) 加藤健太郎, 山田剛史, 川端一光, "R による項目反応理論", オーム社.
(3) 植野真臣, 荘島宏二郎, "学習評価の新潮流", 朝倉書店.
(4) 阿久津洋巳, 石亀雅哉, "項目反応理論を用いた試験問題の検討：共通教育心理学の例", 岩手大学教育学部附属教育実践総合センター研究紀要第11号（2012）, pp. 167-175.
(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) 巨海玄道, 野田常雄, 江藤徹二郎, 中村文彦, "久留米工大における物理学初年次教育の試み", 久留米工大研究報告, No. 38（2015）, pp. 33-41.
(7) 藤田哲也, "初年次教育の目的と実際", リメディアル教育研究, 第1巻第1号（2006）, pp. 1-9.
(8) 大友賢二, "項目応答理論", 電子情報通信学会誌, Vol. 92, No. 12（2009）, pp. 1008-1012.
(9) 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.