

計算力テストを考える

船橋旭高等学校 大橋 真也

1 はじめに

数学会の研究委員会が毎年実施している「計算力テスト」に参加する機会があった。このテストは、新1年生の学力の一部でもある、計算力を測定するために実施されているものである。計算力テストの個々の設問について、項目反応理論の考え方をを用いて、問題の分析を試みようと考えた。

ただしこの分析に関しては、本校の一部の生徒のデータを元にしており、サンプル数に関しても小さいので、正確な分析とまではいかないものの、ある程度の傾向が見えてきたので、ここで報告をする。

2 IRT とは何か

項目反応理論 (IRT, Item Response Theory) とは、1940～50年代以降に発達してきた現代テスト理論の一つであり、古典的テスト理論の様々な問題点を解消するために考えられ、以下のような利点を持つ考え方である。

- 異なるテストにおいても共通の尺度で能力の測定ができる。
- 受験者集団に依存せず、共通の尺度で能力の測定ができる。

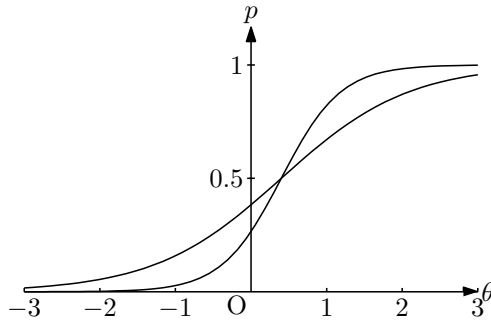
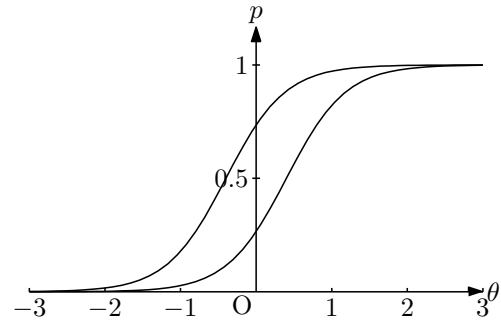
2.1 項目特性曲線

項目特性曲線 (ICC, Item Characteristic Curve) とは、その項目 (テスト問題) の受験者の能力とテスト問題の正答確率を表した曲線であり、ここでは2パラメータロジスティックモデル (2PLM) を例に簡単に説明する。2PLM は、累積正規分布曲線を元にして、以下のような式で表される。

$$p_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}}$$

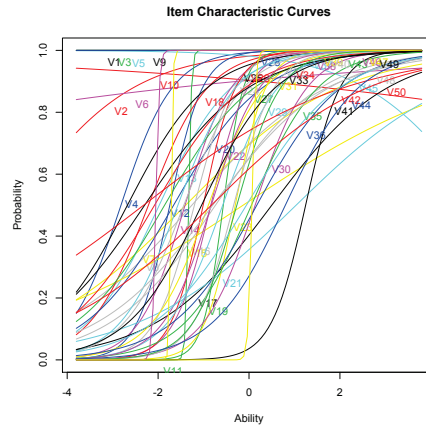
この式で、 D は尺度要素と呼ばれ、累積正規分布曲線に近似するために通常は $D = 1.7$ が用いられている。また a, b がこの曲線のパラメータであり、 a は項目識別度と呼ばれ、曲線の傾きの度合いを表している。つまりその問題の正誤が受験者の実力をどれだけ明確に判別できて

いるかを表すパラメータである。また b は項目困難度を表しており、そのテスト問題の難しさを表している。グラフでは右上がりに増加している位置を表している。それらのパラメータを設定して書いたグラフが以下のグラフである。このグラフでは横軸が受験者の能力を表し、縦軸がその項目に対して正答できる確率を表している。

図 1: $a = 1.5$ と $a = 0.7$ のグラフ図 2: $b = -0.4$ と $b = 0.4$ のグラフ

3 IRT による問題分析

これらの考え方をを用いて、「計算力テスト」全 50 題を 2PLM に当てはめ、ICC を描くと以下のようなになる。



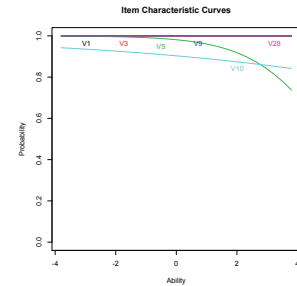
これらのグラフは複雑であるが、曲線の傾向を見ると、いくつかの項目について特徴的なことがわかる。なおグラフにある番号は問題番号順につけた連番である。

3.1 項目識別度による分析

3.1.1 項目識別度が 0 以下の項目

項目識別度は、ICC の傾きを表すため通常は 0 以上であるはずである。しかしサンプル数が少ないことなども関係して、項目識別度が 0 または負の値になったものがあった。それが以下の 6 つの項目 (問題) である。

番号	問題	困難度	識別度
1(1)	$27 + 18 =$	$-2.260467e+11$	0.000
1(3)	$5 - (-2) =$	$-2.260467e+11$	0.000
1(5)	$36 \div (-12) =$	$5.144000e+00$	-0.769
2(1)	$\frac{1}{7} + \frac{3}{7} =$	$-2.260467e+11$	0.000
2(2)	$\frac{3}{5} - \frac{1}{3} =$	$1.512300e+01$	-0.148
7(4)	$-2x \times (-3x)^2 =$	$-2.260467e+11$	0.000



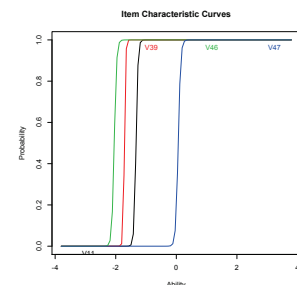
受験者能力に依存しない、あるいは能力を測ることにはならないと考えられるこの問題は、当然のことながら平易すぎる問題であろう。サンプル数が大きくなった場合、状況が変化することもあるかもしれないが、ある程度は納得のいく結果である。

しかし、これらの問題は、受験者の能力を測定するのに意味がないと判断されるが、テストの問題として必要がないわけではない。ある程度誰でもできるような問題が、テストの中になれば、受験生のやる気もなくなってしまうだろう。ただし、これらの問題の出題順序などには考慮が必要であると考えられる。

3.1.2 項目識別度が極端に大きい項目

項目識別度がかなり大きな数値になったものもある。これは、問題の出来不出来が極端であり、受験者の能力を段階的に表すのには不向きである問題などを表していることもある。項目識別度が極端に大きいものは次の4項目ある。

番号	問題	困難度	識別度
2(3)	$\frac{2}{3} \times \frac{1}{3} =$	$-1.330000e+00$	30.808
9(4)	$x^2 - 7x + 10 =$	$-1.714000e+00$	49.032
12(1)	$4x - 5 = 7$	$-2.049000e+00$	25.882
12(2)	$\frac{x}{5} + \frac{7x-5}{10} = 4$	$6.200000e-02$	24.850

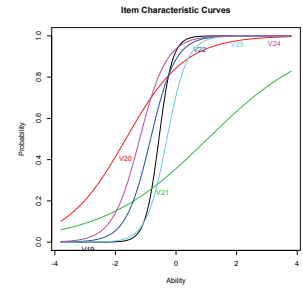


9(4) は因数分解、12 は方程式を解く問題である。上の3題については、困難度も低いが出来不出来がはっきりしている問題である。12(2) については、ある程度平均的な問題ではあるが、識別度が高い問題である。項目識別度が高い問題は、このテストの意味や目的がどのようなものであるのかを考慮して、改善する必要があるかもしれない。この「計算力テスト」が入試のような選別のためのテストならばよいのであるが、受験者の能力判定のためのテストであるならば、検討が必要である。このような出来不出来は、どのような要因から起こるのだろうか。解法などの知識の有無のみで起こっているのだろうか。

3.2 項目困難度による分析

次に項目困難度に着目して、問題を見てみよう。以下は6番のルートの計算に関しての6題について、分析した結果である。

番号	問題	困難度	識別度
6(1)	$\sqrt{6} + \sqrt{12} =$	-5.650000e-01	4.484
6(2)	$\sqrt{27} \div \sqrt{3} =$	-1.658000e+00	1.016
6(3)	$\sqrt{3} \div \sqrt{6} \times (-\sqrt{18}) =$	1.038000e+00	0.570
6(4)	$\sqrt{18} - \sqrt{8} + \sqrt{2} =$	-8.400000e-01	2.440
6(5)	$(2\sqrt{5} + \sqrt{3})(\sqrt{5} - 2\sqrt{3}) =$	-3.020000e-01	2.971
6(6)	$6\sqrt{2} - \frac{8}{\sqrt{2}} =$	-1.201000e+00	2.250



同じルートの計算力を測る問題であっても、困難度の違いと共に識別度の違いも現れている。問題を作成する場合、難易度のみを意識する傾向があるが、識別度を揃えて考える必要もあるだろう。実際には困難度と識別度の2つの要素程度を考慮してテストのデザインをしていくと、受験生の能力を測定できる効果的なテストになると考えられる。

4 おわりに

今回の分析は、少ないサンプルの中での試行的な分析であり、「計算力テスト」に対しての一般的な分析ではない。しかしサンプルが大きくなれば、今後「計算力テスト」だけでなく、様々なテストの分析にこのような考え方が活用できると考えることができる。