

大学生を対象とした金融リテラシーの 項目反応理論に基づく検討

阿 部 圭 司 ・ 小 澤 伸 雄
木 下 康 彦

1. はじめに

近年、わが国では国民一人ひとりが、より自立的で安心かつ豊かな生活を実現するためにも、金融リテラシーを身に付ける必要があるとされている。金融庁は2012年に金融経済教育研究会を設置し、翌2013年にはその報告書において、「生活スキルとして最低限身に付けるべき金融リテラシー」をまとめている。これらは「家計管理」「生活設計」「金融知識及び金融経済事情の理解と適切な金融商品の利用選択」「外部の知見の適切な活用」の4分野から構成されている。また、これを受けて金融経済教育推進会議は「金融リテラシー・マップ」(2014年6月公表、2015年6月改訂)を作成した。これは、小学生から高齢者まで7つの年齢層別に、身に付けるべき金融リテラシーを体系的かつ具体的に示しており、関係当局、業界団体、学校、自治体などが金融リテラシー教育に取り組む際の指針となっている。

同時に、様々な属性を持つ人々を対象とした金融リテラシー水準の調査、金融リテラシー水準と経済・投資行動との関連を調査する試みがみられるようになった。我々もまた、阿部・小澤・木下(2019、2020)を通じて、2度大学生を対象にアンケート調査を行い、大学生の金融リテラシーの水準とその要因について調査、考察を行ってきた。その一方で、金融リテラシー水準を把握するための設問の巧拙、適切さについては、分析の過程で議論することはあっても、具体的な指標を基に検討する機会を持たなかった。また、金融リテラシーに関する先行研究と比較しようにも、設問の内容や難易度の違い、被験者の属性等により、素点や偏差値といった数値で直接比較することは難しいと考えられる。

ところで、計量心理学の領域において、テスト理論の1つとして発展、普及している手法に項目反応理論(Item Response Theory; Item Latent Theory、IRT)がある。これは素点や偏差値などを利用した古典的なテスト理論と比較して、テストを受けた集団や問題の難易度に直接依存せずに被験者の評価をすることが可能とされている。資格試験や能力試験などの開発、評価に用いられているもので、欧米では広く利用されており、日本ではTOEFL、日本語能力試験(JLPT)、ITパスポート試験、SPI2などで採用されている。そこで本稿では阿部・小澤・木下(2019、2020)で得られたサンプルを用

いて、IRTによる分析を試みる。具体的にはパラメータ推計を通じた各設問の評価、他の先行研究との比較から対象となった大学生の金融リテラシーに関する傾向の分析を行い、大学教育におけるヒントを得ること、被験者の能力値に対する推計値と被験者の属性との関連分析などを研究の目的とする。

本稿の構成は以下の通りである。続く第2節では、本稿で採用する分析手法であるIRTとこれを用いた国内外の先行研究をレビューする。第3節では分析対象データと正答率の基本統計、IRTによるパラメータ推計について示す。第4節では推計によって得られた各パラメータの解釈を通じて設問の難易度、識別力、被験者の能力に関する分布及び各種属性との関係について考察する。推計されたパラメータから、全体としては難易度が低くなる傾向が得られた。さらに広い年代を対象とした先行研究と比較すると、難易度は全体と比べると多少低くなることと、いくつかの設問では識別力に大きな差異が認められるものが見つかった。識別力の低さは、金融リテラシー能力の識別という点で問題があり、設問の仕方に検討が必要である指摘することができた。最後の第5節で本稿のまとめと今後の課題について述べるものとする。

2. 項目反応理論と先行研究

2.1. 項目反応理論

項目反応理論では被験者 i が持つ能力を θ_i とする。各設問（以下「項目」と表記）の正答確率が θ_i の関数と考え、項目 j の正答確率を $P_j(\theta_i)$ と表すことにする。被験者が十分に多い場合、 θ_i の分布は正規分布に従うと考えることができる。このとき、被験者の能力値と各項目の正答確率の関係は、正規分布の累積密度関数で表現が可能となる。さらに、操作を簡単にするために正規分布の累積密度関数をロジスティック関数で近似させることを考えると、正答確率は、

$$P_j(\theta_i) = \frac{1}{1 + \exp(-Da_j(\theta_i - b_j))}$$

と書くことができる。 a_j と b_j は項目パラメータと呼ばれ、このモデルは2パラメータ（ロジスティック）モデル（2PLM）と呼ばれる。 a_j は項目識別力パラメータ、 b_j は項目難易度パラメータという。また、 D は標準正規分布の分布関数に近似させるために用意された尺度因子と呼ばれる定数で、1.7が設定されることが多い¹。

この他、項目識別力パラメータ a_j を各項目に共通の定数とし、 b_j だけをパラメータとして推計されたモデルが1パラメータモデル（1PLM）、あるいはRaschモデルと呼ばれる。さらに2PLMを発展させ、選択肢で問われた項目に起こりうる当て推量での正

1 a_j は D とまとめて表記される場合もある。また、パラメータを推計するプログラムによっては a_j と D がまとめて推計されるものも存在する。

解があることを考慮した3パラメータモデル(3PLM)がある。

図1は2パラメータモデルで示される能力値 θ と正答確率の関係を例示している。 b_j は、正答確率が50%になる際の θ の値(水平方向の位置)を意味し、項目の難易度に相当する。図1のcase2($b_j = 0$)はcase1($b_j = -1$)より難しく、case3とは同じであることを示している。一方、項目識別力 a_j は正答確率50%における曲線の傾きであり、受験者の能力を識別する力を表している。傾きが強いほど、ある一定の特性値 θ までは正解者数は少ないが、 θ を越えると急に正解者数が増加することを意味する。 a_j で示す傾きが強い方が、強い識別力を示す。case1とcase2では識別力は同じ($a_j = 1$)であるが、case3は $a_j = 0.3$ と識別力が低い。case3はcase2と比較して能力値 θ が低くても正解者数は多いが、 θ を越える辺りからcase2の正解者数が急激に伸びる、case2の方がより水準 θ を正解・不正解の分かれ目として強く認識できる、という理解が可能だろう。

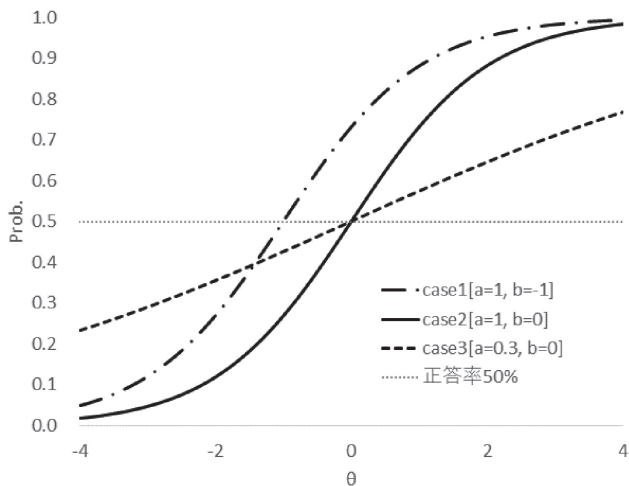


図1：2パラメータモデルによるIRTの例

2.2. IRTを用いた先行研究

IRTを用いた研究事例は様々な分野で見ることができる。大学生に対して行われた試験、アンケート結果をIRTに基づき分析した近年の国内研究事例を挙げると、新入生のコンピュータ・リテラシーについて分析した川合・福山・岩瀬・半田(2010)、一般教養課程における心理学の試験を対象とした阿久津・石亀(2012)、法科大学院適性試験を対象とした荒井・椎名・小牧(2012)、工学系の初年次学生に対し数学の基礎力を問うテストを対象とした中嶋・佐々木(2017)などがある。

金融リテラシーに関する設問を対象とした研究にKnoll and Houts(2012)がある。彼らは米国で行われた3つの調査から集められた18歳以上の6,001名のサンプルを基に、20項目について2パラメータモデルを推計している。推計結果からは債券価格と金利の

関係、ローン金利と返済期間、退職貯蓄制度、年金制度などは難易度が高く、株式投資信託、リスク分散などは識別力が高い、という傾向が報告されている。また、金融知識は性別（男性の方が女性よりも）、年齢、収入の多さ、被験者の金融知識・経済・数学に対する自己評価、貯蓄行動等と正の相関を有していることも報告されている。Kunovskaya, Cude and Alexeev (2014) らは世界銀行による2011年の調査報告から、アゼルバイジャン、ルーマニア、ロシアの3か国の18歳以上、計4,719名のサンプルを基に、6つの項目を用いて1パラメータモデルを推計、国ごとの比較を試みている。分析結果からは、値引きと割引の区別については3か国とも共通して難易度は低く、難易度の高い設問については3か国ともバラバラである、という結果が得られている。また、de Clercq (2019) は2015年にOECDによるINFE (International Network on Financial Education) が実施した30の国と地域の成人を対象とした金融リテラシー調査のうち、11の国と地域から得られた、18歳から79歳までの15,936名のサンプル、7つの項目を用いて1パラメータモデルを推計している。分析結果からは11の国と地域すべてで複利計算が最も難易度が高い項目に、7つの国でローンの利払いが最も難易度の低い項目となったことが報告されている。

国内の研究事例には國方 (2019) がある。金融広報中央委員会が実施した「金融リテラシー調査 (2016年)」の個票データ (25,000名) に対し、2パラメータモデルを推計し、分析を行っている。分析結果として (1) 全般的にやや易しい傾向、(2) 金利と債券価格の関係、トラブル相談窓口に関する設問におけるパラメータ推計値が他の項目と大きく異なっていた、また、Knoll and Houts (2012) による米国の結果を等化したうえで比較し、(3) 複利計算に関する設問では日本の正答確率が米国のそれより低い、(4) インフレや分散投資などでは日米の差はあまり見られないが、住宅ローンなどでは日本の正答確率が高い、という傾向を報告している。

3. 調査対象と分析方法

3.1. 調査対象

本研究で分析対象とするサンプルは、高崎経済大学経済学部の開講科目である「証券論」および「ファイナンシャル・リテラシー」にて、学生が自身のスマートフォン等によりアンケートWebサイトにアクセスし、用意されたアンケートに対して回答する形で得た²。前者は2018年10月15日に、後者は2019年5月1日に実施し、回収後のデータクリーニングにより、有効回答として得られた計501名分である³。分析対象はアンケー

2 アンケートの実施にはREAS (Real-time Evaluation Assistance System) を用いた。詳しくは阿部・小澤・木下 (2019) を参照。

3 アンケートでは、スマートフォン等を自宅に置いてきた、アクセスできなかった等の学生に対しては、同時に紙による質問票を配布し回答を得ている。また、アンケート回収に要した時間 (ログインから回答終了) の短い回答、目視で不正回答と判断できる回答、欠損のある回答等を削除した。

トにおける金融リテラシーの知識を問う22問である。これらは1問を除きすべて選択肢による回答となっており、回答（反応）に対し正解を1、不正解を0とした501×22の行列を作成した。この他、アンケートにより回収した性別、学年、「ファイナンシャル・リテラシー」受講の有無、家庭環境などの要素との関連についても分析し、考察を行う。

3.2. テストの結果とIRTに基づく推計

全22問によるテストの正答率を表1にまとめた。全サンプルでは22問の平均は正答率71.0%、最大94.6%（Q7：インターネット取引での注意）、最小21.0%（Q14：債券価格と金利の関係）であり比較的高い。正答率の低い設問にQ14の他、Q19（46.1%：預金保険制度の理解）、Q20（50.7%：為替（円高・円安）の理解）、Q17（51.7%：保険についての理解）など、大学生にとっては身近ではないテーマが目立つ。

また、表1ではサンプルを被験者の属性で区別し、比較した結果も示している。性別で比較した場合、全体では男性の正答率が高い。大きく正答率に差がある項目もいくつか観察される。各項目の正答率に差異があるかどうかをカイ2乗検定で検証した結果⁴、

表1：項目毎の正答率比較

設問	全体	男性	女性	χ^2	リテラシ 受講者	新入生	χ^2
Q1 適切な収支管理	0.796	0.803	0.785	0.145	0.769	0.845	3.084*
Q2 家計管理クレジットカード	0.780	0.781	0.779	0.000	0.757	0.836	3.194*
Q3 複利と期間についての理解	0.792	0.788	0.801	0.061	0.823	0.753	2.290
Q4 人生の三大費用	0.641	0.628	0.663	0.468	0.651	0.680	0.252
Q5 契約にかかる基本的な姿勢	0.896	0.888	0.912	0.486	0.923	0.895	0.592
Q6 金融トラブルに巻き込まれないための行動	0.872	0.866	0.884	0.204	0.888	0.890	0.000
Q7 インターネット取引での注意	0.946	0.941	0.956	0.267	0.982	0.936	3.815*
Q8 預金金利についての理解	0.892	0.938	0.812	17.607***	0.899	0.890	0.014
Q9 複利についての理解	0.709	0.756	0.624	9.118***	0.787	0.630	10.409***
Q10 インフレと購買力	0.657	0.694	0.591	4.952**	0.645	0.676	0.279
Q11 インフレについての理解	0.820	0.841	0.785	2.102	0.828	0.827	0.000
Q12 住宅ローンを組む際の金利についての理解	0.872	0.878	0.862	0.147	0.911	0.822	5.629**
Q13 資産形成における分散	0.749	0.772	0.707	2.238	0.840	0.662	14.809***
Q14 債券価格と金利の関係	0.210	0.247	0.144	6.827***	0.237	0.160	3.138*
Q15 金利が変化した場合の判断	0.573	0.588	0.547	0.620	0.692	0.498	14.063***
Q16 家族構成の変化に応じた保険の見直し	0.573	0.556	0.602	0.819	0.426	0.699	27.987***
Q17 保険についての理解	0.517	0.559	0.442	5.918**	0.704	0.361	43.653***
Q18 複利（72の法則）についての理解	0.631	0.697	0.514	15.858***	0.651	0.571	2.239
Q19 預金保険制度の理解	0.461	0.478	0.431	0.855	0.468	0.452	0.040
Q20 為替（円高・円安）の理解	0.507	0.566	0.403	11.544***	0.450	0.516	1.422
Q21 金融トラブル回避のための行動	0.884	0.888	0.879	0.025	0.935	0.845	6.709**
Q22 複雑な金融商品を購入する際の適切な行動	0.850	0.828	0.890	2.956*	0.852	0.858	0.001
平均値（正答率）	0.710	0.726	0.682		0.733	0.696	
平均値（正答数）	15.629	15.978	15.011		16.118	15.301	
サンプルサイズ	501	320	181		169	219	

注：p値がそれぞれ ***<0.01、**<0.05、*<0.1であることを示す。

4. 本稿での分析にはすべてR（64bit版、version 3.6.2）および関連するパッケージを用いた。

Q8（預金金利についての理解）、Q9（複利についての理解）、Q10（インフレと購買力）、Q14、Q17、Q18（複利（72の法則）についての理解）、Q20では男性が、Q22（複雑な金融商品を購入する際の適切な行動）のみ女性の正答率が統計的に有意に高い、という結果を得た。

さらに、大学における教育を受けた経験の有無による比較を目的として、ファイナンシャル・リテラシー受講経験のある学生と、新入生（2019年調査の1年生）⁵を比較した。正答率に大きな差があり、有意水準5%以下で有意差が認められた項目は、Q9、Q12（住宅ローンを組む際の金利についての理解）、Q13（資産形成における分散）、Q15（金利が変化した際の判断）、Q16（家族構成の変化に応じた保険の見直し）、Q17、Q21（金融トラブル回避のための行動）である。全体を通じてファイナンシャル・リテラシー受講経験者の方が高い正答率となり、大学における金融教育については一定の効果があると判断できる。

次にIRTの分析を進めるに際して前提となる条件を確認する。IRTでは受験者が持つ能力 θ のみで正答率が決定するとしているため、テストにおける項目で測定される能力は1つだけと仮定している。これを次元性と呼ぶ。次元性を確認するために、因子分析やスクリープロットが推奨されている。本稿でもスクリープロットを図2のように作成した。図は横軸が因子の数、縦軸が固有値である。

図から、第1因子の固有値が3.345と特に高く、第2因子（1.658）以降は第8因子（1.018）まで1を超える固有値を持つものの、水準は第1因子と比較して十分に低い。このことから次元性は満たされていると判断する。

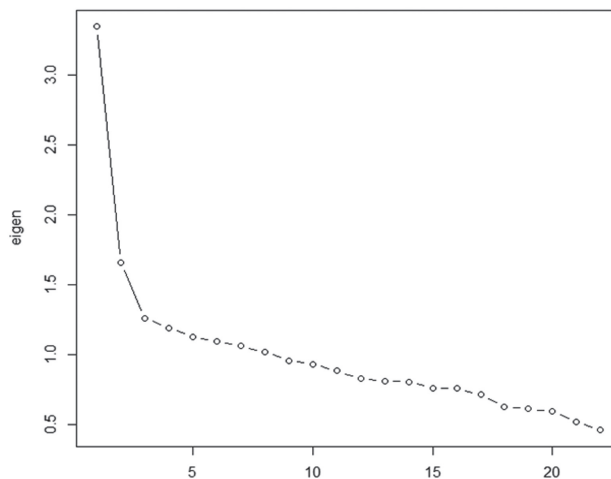


図2：金融リテラシーテストにおけるスクリープロット

5 2018年調査における1年生は調査時点が10月であり、ファイナンシャル・リテラシーをはじめとした大学の講義を半期受講しているため、新入生のサンプルから除いている。

評価に用いるモデルとしては、1 から 3 パラメータモデルをそれぞれ周辺最尤推定によりパラメータを推計し⁶、AIC (Akaike's Information Criteria) の水準に基づき 2 パラメータモデルを選択した⁷。

4. 分析結果とその検討

4.1. 推計されたパラメータ及びICCによる各項目の評価

表 2 に推計された識別力 a_j 、難易度 b_j の各パラメータ及び標準誤差を示す。識別力は尺度因子が乗じられたものを示している。また、各項目の 2 パラメータの水準を理解しやすくする目的でこれを散布図で示したものが図 3 である。横軸は難易度、縦軸は尺度因子×識別力を示し、後述する ICC の特性で分類した 5 グループで分けて表記している。

まず、難易度のパラメータは-4.750から9.746の範囲を取り、平均-0.971であるが、表より難易度が 0 を超える項目は 2 つしかない。最大値の 9.746 (Q14) を除くと次に大きな値は 0.438 (Q19) となり、平均値は-1.482 となることから全体的な難易度は低いと判断できる。図 3 からほとんどの項目が 0 より左に位置していることから、難易度が低い傾向にあることが確認できる。推計精度を標準誤差から確認すると、難易度が極めて高く推計された Q14 は 9.656 とバラツキが多い。Q4、Q11 など 1 を超える標準誤差を示しており、これらの項目については推計精度が低いことが分かる。識別力は 0.137 から 2.422 の範囲を取り、平均は 0.916 となった。識別力に関しては、標準的な範囲に収まる傾向にあると判断できる。標準誤差は難易度の標準誤差に比べて低く、推計精度は高いと評価できる。図 3 を見ると、難易度の高い Q14 が他の項目と異なる傾向にあることが

表 2：IRT による 2 パラメータロジスティックモデルの推計結果

項目	1.7 × a (識別力)		b (難易度)		項目	1.7 × a (識別力)		b (難易度)	
	推計値	標準誤差	推計値	標準誤差		推計値	標準誤差	推計値	標準誤差
Q1	0.837	0.165	-1.853	0.322	Q13	0.693	0.145	-1.732	0.343
Q2	1.173	0.191	-1.355	0.185	Q14	0.137	0.136	9.746	9.656
Q3	1.090	0.181	-1.500	0.211	Q15	0.759	0.140	-0.435	0.149
Q4	0.250	0.116	-2.346	1.123	Q16	0.370	0.117	-0.818	0.351
Q5	2.422	0.456	-1.558	0.146	Q17	0.423	0.119	-0.167	0.225
Q6	1.481	0.252	-1.750	0.207	Q18	1.061	0.168	-0.618	0.126
Q7	1.666	0.331	-2.366	0.294	Q19	0.368	0.118	0.438	0.283
Q8	1.095	0.212	-2.316	0.351	Q20	0.339	0.115	-0.084	0.272
Q9	0.698	0.141	-1.404	0.282	Q21	1.665	0.280	-1.732	0.190
Q10	1.113	0.172	-0.724	0.129	Q22	1.017	0.189	-2.024	0.307
Q11	0.327	0.144	-4.750	2.035	平均値	0.916		-0.971	
Q12	1.173	0.210	-2.019	0.278	最大値	2.422		9.746	
					最小値	0.137		-4.750	

6 推計の具体的な考え方、算出の流れについては豊田 (2012)、加藤・山田・川端 (2014) 等を参照。

7 推計により得られた AIC は 1、2、3 パラメータモデルの順で 11, 175.01、11, 055.70、11, 071.91 となった。

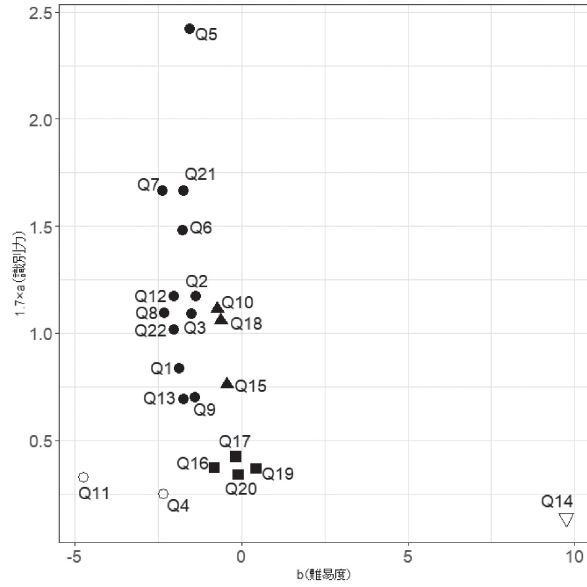


図3：推計された各項目パラメータの傾向

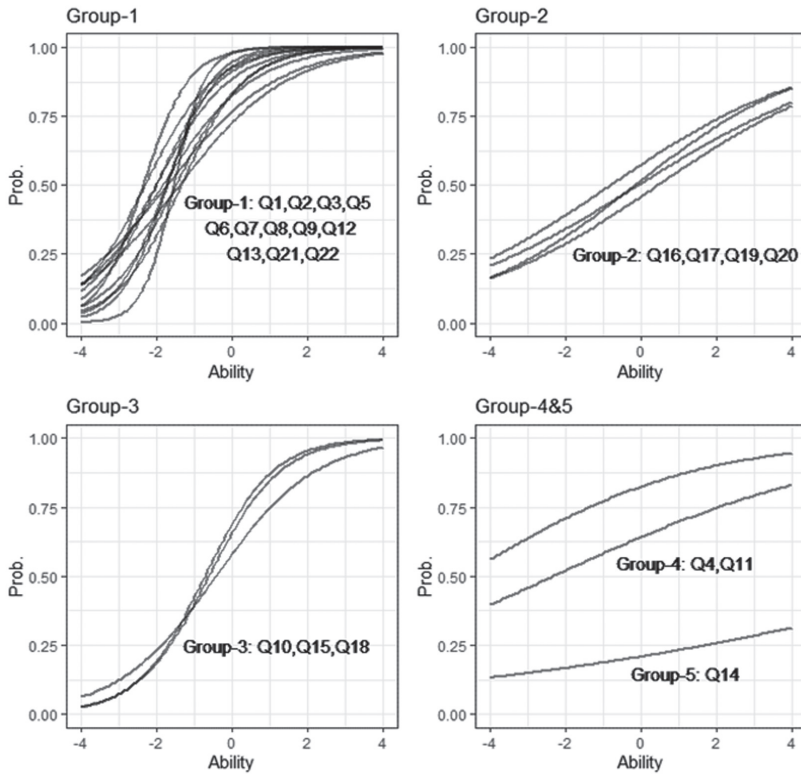


図4：各項目の推計パラメータを基にしたICCとその類型

分かる。識別力の高いQ5も異なるように見えるが、1.7倍されていること、縦軸と横軸のスケールの違いを考慮すると、それほど離れた水準ではなく、むしろQ14に加えてQ11が異なる水準にあると考えられる。

次に、各項目の推計された2パラメータを基に描かれるICC (Item Characteristic Curve: 項目特性曲線) を図4に示す。図4では主観的ではあるが、ICCを大きく5つグループに分類し、検討を試みる⁸。第1のグループは最も多い12項目が分類されているケースで、識別力は標準的(平均値1.301)であるが、全体的に難易度は低い(同-1.837)項目群である。第2のグループは4項目が分類されている。難易度は平均的(-0.158)であるが、識別力が低い(0.375)である項目群である。第3のグループは第1のグループに近いが、難易度はそれらよりも平均に近い(-0.592)項目群である。第4のグループはQ4(人生の三大費用)とQ11(インフレについての理解)からなり、難易度(-2.833)、識別力ともに低い(0.425)項目群、最後の第5のグループはQ14(債券価格と金利の関係)のみで難易度が極めて高く(9.746)、識別力が低い(0.137)ものである。第4、第5のグループは標準誤差も大きく、推計の精度も低い。第4のグループに入る2つの設問は金融リテラシーにおいても重要な要素であるだけに、設問の仕方が適切ではなかったと判断することができる。また、Q14は債券価格の式を理解していることを求める設問で、やや専門性が高く、多くの調査で正答率が低いと報告されている設問である。金融リテラシーに求められる内容であることは否定しないが、現状を踏まえた設問を、また金融リテラシー教育の中では、その効果的な教授法について考慮しなくてはならない項目であろう。

さらに、國方(2019)で推計されたパラメータを規準集団とし、今回推計されたパラメータを等化し、比較を試みる⁹。我々の項目は國方(2019)で分析対象とされた金融リテラシー調査(2016)に準拠しているため、22問中21問が共通項目となる。等化に必要な等化係数についてはMean & Mean法を用いた。図5は國方(2019)で得られたICCと今回の推計パラメータを等化し、得られたICCを並べて描いたものである。破線は今回の推計によるもの、実線は國方(2019)で報告されているものである。

多くの項目で2つのICCは同じ傾向を持ち、そのうちいくつかについては破線、すなわち本研究の方で難易度がやや低い、つまり大学生の正答率が高い傾向がみられる。これらは正答率に基づく観察(阿部・小澤・木下(2019))と整合的である。

2つのICCが異なる傾向を示す項目をいくつかまとめると、(1)Q4、Q11(グループ4)は大学生のサンプルから得られたICCは難易度がかなり小さくなっているが、識別力が落ちている、(2)Q16、Q17、Q19(グループ2)では難易度はほぼ同じであるが、識別力が落ちている、(3)Q14(グループ5、債券価格と金利の関係)では難易度、識別力共

8 グループ数の5は任意に定めたが、分類自体はクラスタ数5としたk-mean法によるクラスターリングによる結果を用いている。

9 等化とは異なる版のテスト、異なる被験者から得られた項目パラメータや能力値を相互比較するための手続き。等化の考え方、等化係数の具体的な算出については豊田(2012)、加藤・山田・川端(2014)等を参照。

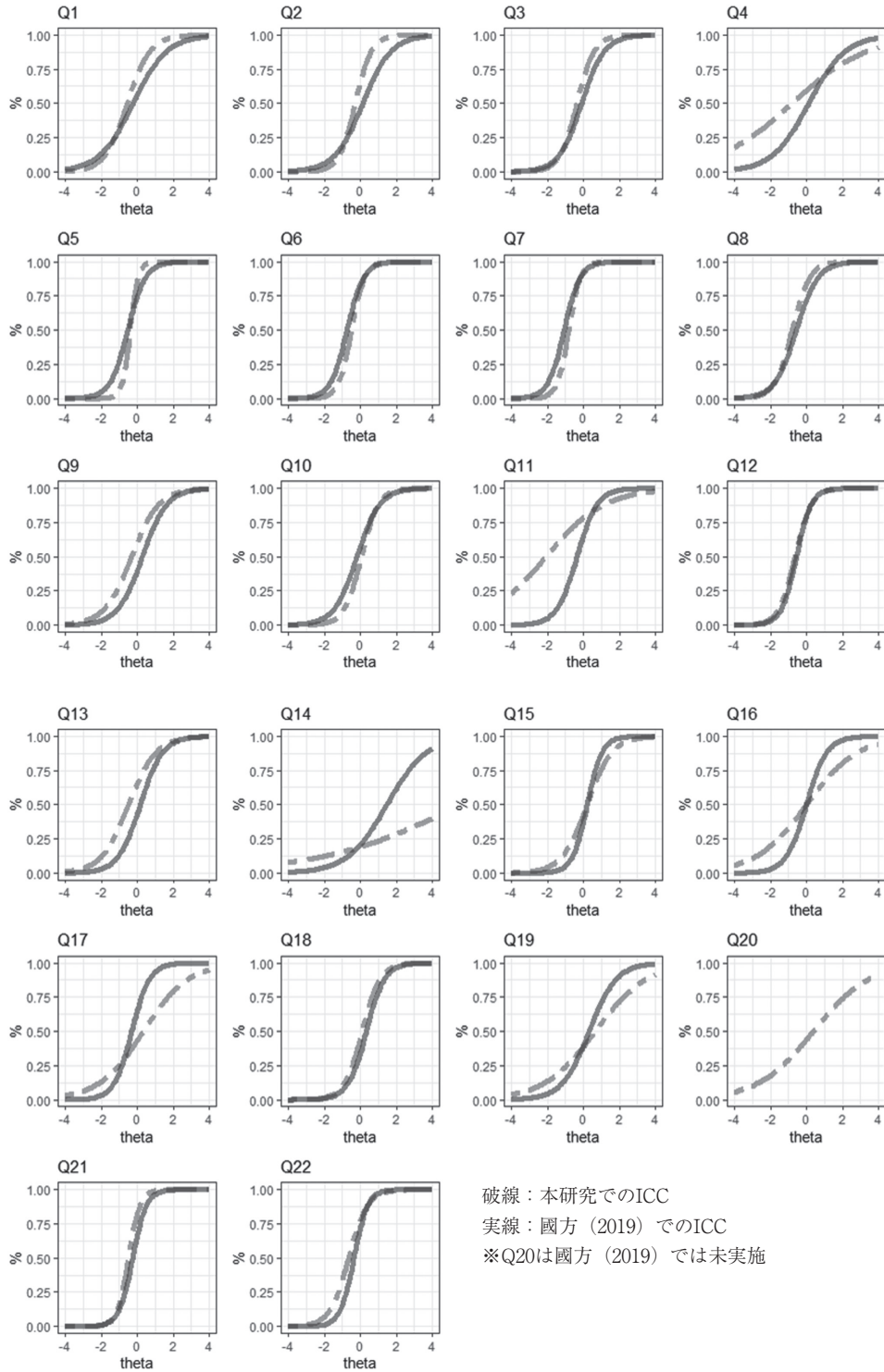


図5：国方(2019)と本研究で得られたICCの比較

に大きな相違が認められる、というものである。グループ4は人生の三大費用（Q4）とインフレについての理解（Q11）であり、グループ2は家族構成の変化に応じた保険の見直し（Q16）、保険についての理解（Q17）、預金保険制度の理解（Q19）である。これらの項目はファイナンシャル・リテラシー、すなわち大学における金融教育の中で扱っている内容である。知識として頭の中には入り易いが、当事者とならないと実感しづらいただろうと我々は評価しており、その影響が識別力の低さという結果に表れている可能性はあるが、要因についてはさらに調査・考察が必要になるものと思われる。また、Q14の傾向は内容が難しい点と、（おそらく被験者全員が）債券投資の当事者ではなく、関心が低いことが影響しているものと思われる。

4.2. 被験者の能力値 θ の傾向と要因分析

IRTでは推計により得られた項目の難易度、識別力と共にこれらパラメータを利用して、被験者の能力値 θ についての推計値を得ることができる。今回の分析サンプル501名及び属性別の θ 推計値の概要を表3に示す。また、横軸に正答数、縦軸に推計された能力値を取り、両者の関係を散布図として示したものが図5である。

表より全体の能力値 θ の平均値は-0.068、中央値は0.019となり、ほぼゼロ付近に中心があることが分かる。属性別に見た能力値 θ の傾向では、男性の平均が0.001に対し、

表3：被験者の能力値 θ

パネルA：全体及び性別の能力値 θ

性別	サンプル サイズ	平均値	標準偏差	中央値
全体	501	-0.068	0.800	0.019
男性	320	0.001	0.820	0.146
女性	181	-0.190	0.751	-0.080
t 値		2.647		
p 値		0.008		
cohen's d		0.241		
検定力（5%）		0.734		

パネルB：学年別の能力値 θ

学年	サンプル サイズ	平均値	標準偏差	中央値
1 年	251	-0.140	0.812	-0.064
（うち新入生）	219	-0.138	0.818	-0.054
2 年	168	-0.020	0.793	0.137
3 年	69	0.013	0.794	0.013
4 年	13	0.282	0.562	0.531
F 値		1.944		
p 値		0.122		
cohen's f		0.109		
検定力（5%）		0.833		

女性が-0.190と1%の有意水準で差が認められる。男性の方が高いという傾向は正答率で分析した阿部・小澤・木下（2019、2020）と同じ傾向である。また、学年別ではサンプルサイズに偏りがあるものの、学年が上がるにつれ、能力値は高くなる傾向が認められる。ただし、1から4年生までの4群間に有意な差異は認められなかった。

図5からは、能力値と正解数には正の相関が存在するが、同じ得点でも能力値は幅を有していることが分かる。同じ正解数でも能力値で計測すると、ばらつきが存在している。問題により難易度が異なるため、同じ正解数でも能力値の評価が異なるためである。また、図中の実線は能力値（Theta）に対し正解数（Score）を回帰させたときの回帰直線である¹⁰。概ね回帰直線が分布の中心を貫いているが、能力値・正解数の低い部分では回帰直線は分布の下側にあることが特徴である。

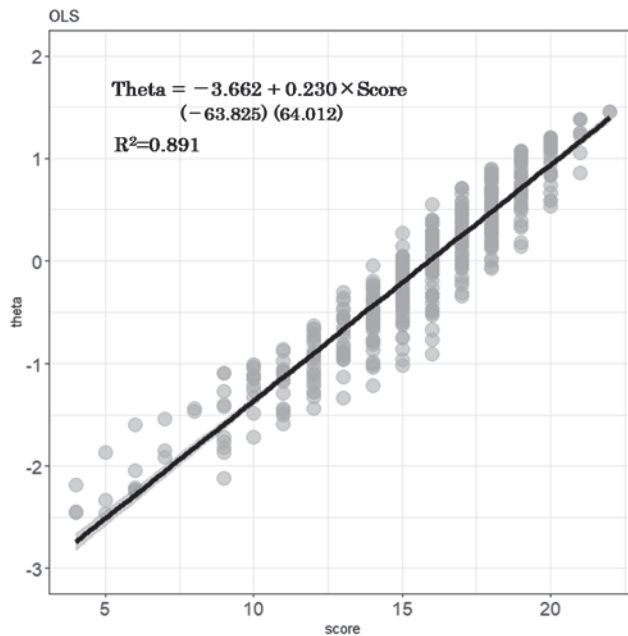


図6：正答数と能力値 θ の相関

次に能力値の推計値を応答変数、能力値に影響を与えると考えられる要因を説明変数として要因分析を試みる。この分析も応答変数に正答数を用いた阿部・小澤・木下（2019、2020）と同じアプローチではあるが、図5のように同じ正答数でも能力値に差異があることから、新たな知見が得られると考えた。推計方法としては回帰分析を用いた。能力値に影響を与える説明変数には、経済学部で開講されている「ファイナンシャル・リテラシー」の受講の有無（有り=1、無し=0）、高校までに金融教育を受けた機会（有り

10 図6中の回帰式下のかっこ内の数値はt値を示している。

= 1、無し=0)、保護者から教わる機会（有り=1、無し=0）、家族に金融機関に勤務する者（有り=1、無し=0）、性別（男性=1、女性=0）、一人暮らしかどうか（一人暮らし=1、家族と=0）、経済・金融情報へのアクセス頻度（5件法による選択肢を区別するため「まったくない」をゼロとした4つのダミー変数）を用いている。推計結果を表8に示す。

正答数と能力値の間に高い正の相関があることから想像できるように、能力値に関係している要素は我々の先行研究で得られた3つの結果と同じ傾向を有している。1つは、ファイナンシャル・リテラシー受講者、保護者から教わる機会がある者ほど能力値は高くなる、という金融教育受講の機会が貢献している傾向である。2つ目は経済・金融情報へのアクセス頻度が高い者ほど能力値が高くなる、という自主的な経済・金融知識獲得の行動による傾向である。金融教育受講の機会については、ファイナンシャル・リテラシー受講の要素は5%水準で有意となり、先行研究よりも強い結果が得られた。また、同じ金融教育受講であっても、高校までに金融教育を受講した経験と、家族に金融機関に勤める者がいる要素は先行研究と同じく有意な反応は得られなかった。さらに、3つ目として男性の方がより能力値が高くなる、という傾向が得られた。性差については今回の分析でも明確な原因は明らかにできていない。性差が直接の要因となる可能性もあるが、何らかの代理変数として表れている可能性もあり、引き続き調査を行う必要性があるといえる。

表4：被験者の能力値と環境変数との関連

OLS 説明変数	応答変数：theta（能力値）		
	係数	標準誤差	p 値
切片	-0.496	0.117	0.000***
ファイナンシャル・リテラシー受講の有無	0.151	0.074	0.042**
高校までに金融教育を受けた機会	-0.091	0.079	0.251
保護者から教わる機会	0.188	0.073	0.010*
家族に金融機関に勤務する者	-0.058	0.087	0.509
性別	0.156	0.074	0.035**
一人暮らし	0.022	0.080	0.779
情報へのアクセス頻度（月1より少ない）	-0.024	0.131	0.856
情報へのアクセス頻度（月1回程度）	0.176	0.131	0.180
情報へのアクセス頻度（週1回程度）	0.345	0.106	0.001***
情報へのアクセス頻度（ほぼ毎日）	0.212	0.115	0.067*
自由度調整済み決定係数	0.056		
F 値（p 値）	3.977 (0.000)		
Cohen's f^2 / 標本検定力（5%）	0.060 / 0.944		
AIC	1182.68		
サンプルサイズ	501		

※ ***<1%, **<5%, *<10%

5. おわりに

本稿では大学生に対して実施した金融リテラシー調査に基づき、項目反応理論による項目（設問）の評価、能力値の要因分析を試みた。主だった結果は次の通りである。

推計された難易度のパラメータから、全体的に難易度は低い。難易度の低さは高い正答率と整合的な結果である。項目反応理論では難易度に加えて、識別力のパラメータによる視点で各項目を評価できる点がメリットと考えられる。推計された2つのパラメータの傾向から全22問の反応をいくつかの型に分類することができた。さらに、幅広い年代層を対象とした「金融リテラシー調査（2016）」の個票データを用いた國方（2019）の推計結果を用いて等化し、比較すると、難易度の低さを別にすれば各項目の反応は概ね一致する傾向にあると評価できるが、いくつか一般的な傾向とは異なる項目が見つかった。例えば、Q14（債券価格と金利の関係）は難易度が高く、識別力が低い。先行研究においても正答率が低く、國方（2019）による同じ項目の推計結果と大きく異なる傾向を有している。債券価格の一般式における金利の効果という設問自体の難易度以外にも、債券が大学生にとっては馴染みのない金融資産であることが影響していると推測できる。債券という金融資産とその特徴をどれだけ認識しているか、という設問を入れるなど、段階的な調査設計が必要であろう。

次にQ4（人生の三大費用）、Q11（インフレについての理解）は逆に難易度が低く、識別力も低い。また、Q16（家族構成の変化に応じた保険の見直し）、Q17（保険についての理解）、Q19（預金保険制度の理解）、Q20（為替の理解）は平均的な難易度だが、識別力が低く、項目特性曲線（ICC）は一般よりもフラットな形状を構成している。能力値に関わらず一定割合の正解者が存在することを示唆しているため、金融教育の効果ともいえるが、逆に当て推量による結果とも考えられる。いずれにせよ、Q14も含めてこれらの項目は金融リテラシーの能力を識別しているとはいえず、出題の仕方を改良し、ICCが改善されるかを検証する必要があるだろう。

最後に能力値の背景にある要素を検討するために被験者の属性値との関連を分析した。分析結果からは金融教育と能力値との間に正の関係が、また経済・金融情報へのアクセス頻度という、自主的な情報収集の姿勢が同じく正の関係を有することが確認された。今後は今回の分析を踏まえてアンケートでの設問の仕方を改善し、金融リテラシー水準の適切な計測を目指すと共に、金融教育を通じた金融リテラシー水準改善の分析、各種属性との関連性の分析をさらに進め、これらを通じてより良い金融教育のプログラムを検討する、といったPDCAサイクルの確立を目指すことが求められるだろう。

（あべ けいじ・高崎経済大学経済学部教授）

（おざわ のぶお・高崎経済大学経済学部非常勤講師・CFP）

（きのした やすひこ・株カナザワ・CFP）

参考文献

- 阿久津洋巳・石亀雅哉（2012）「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』、No.11、pp.167-175.
- 阿部圭司・小澤伸雄・木下康彦（2019）「高崎経済大学学生の金融リテラシーに関する研究－金融リテラシー調査（2016）に準じた調査結果－」、『高崎経済大学論集』、Vol.62、No.2、pp.1-21.
- 阿部圭司・小澤伸雄・木下康彦（2020）「大学生に対する金融リテラシーに関する研究－新入生と在学生との比較を中心として－」、『高崎経済大学論集』、Vol.62、No.3・4、pp.1-18.
- 荒井清佳・椎名久美子・小牧研一郎（2012）「項目反応理論を用いた法科大学院適性試験の年度間の比較」、『大学入試センター研究紀要』、No.41、pp.1-14.
- 加藤健太郎・山田剛史・川端一光（2014）『Rによる項目反応理論』、オーム社
- 川合治男・福山裕宣・岩瀬弘和・半田勝久（2010）「項目反応理論による新入生のコンピュータ・リテラシーの測定」、『東京成徳大学研究紀要』、No.17、pp.35-47.
- 國方明（2019）「反応項目理論による金融リテラシー再検証」、日本ファイナンス学会第1回秋季研究大会金融経済教育推進会議（2014、2015）「金融リテラシー・マップ」、金融広報中央委員会
- 豊田秀樹（2012）『項目反応理論[入門編]第2版』、朝倉書店
- 中嶋康博・佐々木良勝（2017）「項目反応理論による数学基礎力確認テストの解析」、『久留米工業大学研究報告』、No.36、pp.63-71.
- De Clercq, B.（2019）“A Comparative Analysis of The OECD/INFE Financial Knowledge Assessment Using The Rasch Model”, Empirical Research in Vocational Education and Training, Vol.11, pp.1-29.
- Knoll, M. A. Z. and C. R. Houts（2012）“The Financial Knowledge Scale: An Application of Item Response Theory to the Assessment of Financial Literacy”, The Journal of Consumer Affairs, Vol.46, No.3, pp.381-410.
- Kunovskaya, I. A., B. J. Cude and N. Alexeev（2014）“Evaluation of a Financial Literacy Test Using Classical Test Theory and Item Response Theory”, Journal of Family and Economic Issues, Vol.35, No.4, pp.516-531.
- Rizopoulos, D.（2006）“ltm: An R Package for Latent Variable Modeling and Item Response Theory Analyses”, Journal of Statistical Software, Vol.17, No.5, pp.1-25.

