

「金融リテラシー調査(2019年)」に基づく項目パラメータの推定 (Estimation of Item Parameters on Financial Literacy: New Survey Evidence)

國方 明*

1. イントロダクション

本稿では、「金融リテラシー調査(2019年)」項目パラメータの推定結果を報告する。

國方(2020)は「金融リテラシー調査(2016年)」個票データに項目反応理論(item response theory, 以下IRTと略す。)を適用して、客観問題25問の項目パラメータを推定した。項目パラメータは、各問の難易度や識別力などの特徴を表すパラメータである。國方(2020)の推定結果を利用する学術研究が散見される。まず、阿部他(2020)は、大学生を対象にした独自アンケートの回答から推定した項目パラメータと、國方(2020)の推定値を比較した¹⁾。次に、國方(2021)は、國方(2020)の項目パラメータ推定値と「金融リテラシー調査(2019年)」個票データとを用いて、2019年における各被験者の能力パラメータを推定した。このように項目パラメータ推定値の公表は有用と考えられる。

2. 分析枠組み

2.1. 「金融リテラシー調査」の紹介

「金融リテラシー調査」は、わが国18歳以上の個人の金融リテラシーの現状把握を目的とするアンケート調査である。本稿執筆時点(2022年8月)で、「金融リテラシー調査」は、2016年、2019年と2022年の3回実施されている。各調査の結果は、金融広報中央委員会(2016, 2019, 2022)で報告されている。

3回の調査で共通して、客観問題25問が出題されている²⁾。正答率は、2016年調査で55.6%、2019年調査で56.6%、2022年調査で55.7%である。

2.2. IRT

IRTでは、被験者の回答パターンにロジスティック関数を当てはめる。本稿では、國方(2020)と同じ2パラメータロジスティックモデルを採用する。2パラメータロジスティックモデルのロジスティック関数は(1)式である。

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

$P_j(\theta_i)$ は、被験者*i*が項目*j*に正答する確率を表す。 θ_i は被験者*i*の能力を表す。 θ_i は $-\infty$ から $+\infty$ の値を取り、 θ_i が大きいほど能力が高い被験者を表す。通常、 θ_i が平均0標準偏差1の標準正規分布に従うと仮定する。横軸に θ_i を測り縦軸に $P_j(\theta_i)$ を測る平面に、(1)式を描く。すると通常右上がりのS字型曲線になる。この曲線を項目特性曲線と呼ぶ。

また、(1)式右辺にある a_j と b_j をまとめて、項目*j*の項目パラメータと呼ぶ。 a_j を項目識別力と呼び、 b_j を項目困難度と呼ぶ。

項目困難度 b_j は、正答確率が50%になるときの θ_i の値に当たる。 b_j も $-\infty$ から $+\infty$ の値を取る。 b_j が小さくなるほど、能力 θ_i の低い被験者でも項目*j*に正答しやすくなる。そこで、項目*j*は、相対的に簡単な項目と判定される。

一方、項目識別力 a_j は、 θ_i と b_j とが等しくなるときの、項目特性曲線の傾きを表す。通常、

* 青森公立大学教授

項目特性曲線は右上がりなので、 a_j は正值になる。 a_j が大きくなるほど、正答確率50%近辺の領域で、項目に正答した人の能力が b_j よりも高く、項目に誤答した人の能力が b_j よりも低いと判断しやすい。

本稿では、次の2つの分析を実施する。

第1に、「金融リテラシー調査(2019年)」個票データを用いて、客観問題25項目にかかわる項目パラメータを推定する。不正解の被験者と「わからない」を選ぶ被験者とを区別しない。項目パラメータの推定方法は周辺最尤推定法である。R(R Core Team, 2022)のirtosysパッケージ(Partchev, 2022)上でIRT Command Language(Hason, 2002)を操作して、項目パラメータを推定した。Rのバージョンは4.2.1、irtosysパッケージのバージョンは0.2.2、IRT Command Languageのバージョンは0.020301である。

第2に、本稿と國方(2020)との間で、項目パラメータ推定値を比べる。比較に際して、國方(2020)の項目パラメータの原点と尺度を基準にして、本稿の項目パラメータの原点と尺度を國方(2020)のそれに揃える。この手続きを等化と言う。等化を実行する方法として、mean & sigma法、mean & mean法、Haebara法、Stocking & Lord法の4つを採用する³⁾。R 4.2.1のplinkパッケージ(Weeks, 2010, 2017)を用いて等化を実施した。plinkパッケージのバージョンは1.5-1であり、本稿と國方(2020)とで同じである。

3. 分析結果の紹介

3.1. 等化前の項目パラメータ

表1の第1列と第2列に、「金融リテラシー調査(2019年)」個票データを用いて推定した項目パラメータを示す⁴⁾。また、表1の第3列と第4列に、國方(2020)で報告した項目パラメータ推定値を参考までに示す⁵⁾。

表1第1列で、 a_j は全て正值である。また、表1第2列で、 b_j は15個の項目で負値になり、10個の項目で正值になる。さらに項目15と項目25で、推定値が他の項目と大きく異なる。以上の

結果は國方(2020)と同じである。

表1下段に、全項目の推定値を使って計算した平均と標準偏差を示す。2019年調査では、項目識別力の平均と項目困難度の平均のどちらも若干小さい。

3.2. 等化後の項目パラメータとテスト特性曲線

表2に、等化を行った後の項目パラメータ推定値を示す。表2の推定値は、表1の推定値とあまり異ならない。

図1に、表1と表2の項目パラメータを使って描いたテスト特性曲線を示す。テスト特性曲線は、横軸に θ_i を測り、縦軸に $P_j(\theta_i)$ の全項目合計を測る平面に、両者の関係を示す曲線である。全体として、6本のテスト特性曲線はほぼ重なっている。但し、次の2つが異なる。まず、横軸が-2以上0以下近辺の領域で、2019年調査(等化前)の線が他の線よりも高い。次に、横軸が0以上2以下近辺の領域で、mean & sigma法の線が、他の線よりも低い。

4. 結果の考察

分析結果は次の2つである。(A)全体として、項目パラメータの推定結果はあまり異ならない。(B)但し、等化を行う前の項目パラメータ推定値と、mean & sigma法で等化を行った項目パラメータの推定値が、他の推定値と若干異なる。

以上の分析結果について、2つの考察を行う。

第1に、2.1節で述べた通り、2019年調査の正答率は、2016年調査のそれよりも高かった。図1では、横軸の能力が平均0以下になる領域で、2019年調査(等化前)の線の高さが他の線よりも高い。そこで、正答率の上昇は、主に能力が平均以下の被験者で生じたかと推測できる。

第2に、表2でmean & sigma法にかかわる項目識別力の平均が小さい。また図1の一部領域で、mean & sigma法の線が他の線よりも低い。これらの結果について、mean & sigma法の一次変換では、2016年調査における項目困難度の標準偏差と2019年調査における項目困難度の標準偏

差との比を用いる。一方、表1では、項目25の困難度が、2019年調査では2016年度調査よりも大きい。そこで、項目困難度の標準偏差が2019年調査で2016年調査よりも小さくなり、mean & sigma法の結果と他の結果とが異なっただと考える。

(2022年5月31日受付)

5. 謝 辞

分析にあたり、金融広報中央委員会（事務局：日本銀行情報サービス局内）から「金融リテラシー調査(2016年)」と「金融リテラシー調査(2019年)」の個票データの貸与を受けた。また、匿名の本誌査読者から丁寧なコメントを頂いた。記して感謝申し上げる。なお、言うまでもなく残された問題は筆者に属する。

- 1) 阿部他(2020)は、日本ファイナンス学会第1回秋季研究大会(2019年11月)で筆者が國方(2020)の草稿を報告した際の推定結果を参照している。そして、報告時の推定結果は、國方(2020)の推定結果と同じである。
- 2) 各項目の設問内容については、國方(2020, 表1)を参照してほしい。
- 3) 各方法の詳細については、加藤他(2014, 第11章と第14章), Kolen & Brennan(2014, Chapter 6), González & Wiberg(2017, Chapter 5)を参照してほしい。
- 4) IRTには次元性と局所独立性という2つの前提が両方成り立たなければいけない。本稿執筆過程でこれら前提が成り立つか否かを検証して、IRTの前提が成り立たないとは言えないと判断した。検証結果がご入り用の方は筆者へのご連絡をお願いしたい。
- 5) 本稿と國方(2020)との間で、Rのバージョンとirtoysパッケージのバージョンが異なる。バージョンアップの影響を評価するために、本稿執筆に際して、國方(2020)で推定した項目パラメータを、R 4.2.1とirtoysパッケージ

0.2.2を用いて再度推定した。今回再推定した項目パラメータの値は、國方(2020)の推定値と全く変わらなかった。なお、IRT Command Languageのバージョンは、本稿と國方(2020)との間で同じである。

引用文献リスト

- 阿部 圭司・小澤 伸雄・木下 康彦(2020). 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 高崎経済大学論集, 63(1), 1-15. <http://doi.org/10.20635/00001083>
- González, J. & Wiberg, M. (2017). *Applying Test Equating Methods Using R*. Springer
- Hason, B. P. (2002). IRT Command Language. Retrieved August 5, 2022, from <https://sourceforge.net/projects/ssm/>
- 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
- 金融広報中央委員会(2016). 「金融リテラシー調査」の結果 Retrieved May 31, 2022, from https://www.shiruporuto.jp/public/document/container/literacy_chosa/2016/pdf/16literacy.pdf
- 金融広報中央委員会(2019). 「金融リテラシー調査2019年」の結果 Retrieved May 31, 2022, from https://www.shiruporuto.jp/public/document/container/literacy_chosa/2019/pdf/19literacy.pdf
- 金融広報中央委員会(2022). 「金融リテラシー調査2022年」の結果 Retrieved August 3, 2022, from https://www.shiruporuto.jp/public/document/container/literacy_chosa/2022/pdf/22literacy.pdf
- Kolen, M. J. & Brennan, R. L. (2014). *Test Equating, Scaling, and Linking Methods and Practices* (3rd ed.). Springer
- 國方 明(2020). 項目反応理論による金融リテラシー再検証 青森公立大学論纂, 5(1・2), 15-28
- 國方 明(2021). わが国金融リテラシー説明要因の再検証:「金融リテラシー調査(2019年)」

の利用 東北経済学会誌, 74 (1), 1-28.
Retrieved August 5, 2022, from <https://sites.google.com/site/touhokukeizaigakkai2/touhoku-keizai-gakkai-shi>

Partchev, I. (2022). irtoys: A Collection of Functions Related to Item Response Theory (IRT). R package version 0.2.2. Retrieved August 5, 2022, from <https://cran.r-project.org/web/packages/irtoys/>

R Core Team (2022). R: A Language and Environment for Statistical Computing. R Foundation for

Statistical Computing. Vienna, Austria. . Retrieved August 4, 2022, from <https://www.R-project.org/>

Weeks, J. P. (2010). plink: An R Package for Linking Mixed-Format Tests Using IRT-Based Methods. *Journal of Statistical Software*, 35(12), 1-33. <https://doi.org/10.18637/jss.v035.i12>

Weeks, J. P. (2017). plink: IRT separate calibration linking methods. R package version 1.5-1. Retrieved August 5, 2022, from <https://cran.r-project.org/web/packages/plink/>

[表1] 項目パラメータ推定結果（等化前）

項目番号	2019年調査		参考：2016年調査	
	項目識別力推定値	項目困難度推定値	項目識別力推定値	項目困難度推定値
1	1.019	-0.236	1.067	-0.228
2	1.253	0.035	1.274	0.141
3	1.648	-0.126	1.586	-0.097
4	0.921	0.140	0.983	0.124
5	1.946	-0.588	2.039	-0.521
6	2.061	-0.772	2.192	-0.736
7	2.267	-1.125	2.252	-1.067
8	1.606	-0.688	1.671	-0.557
9	1.443	0.261	1.468	0.288
10	1.615	-0.163	1.579	-0.185
11	1.760	-0.396	1.798	-0.353
12	2.318	-0.638	2.417	-0.560
13	2.567	-0.861	2.590	-0.783
14	1.616	0.119	1.612	0.172
15	0.942	1.450	0.909	1.476
16	2.374	0.151	2.412	0.214
17	1.699	0.048	1.826	0.135
18	1.687	-0.084	1.765	-0.002
19	1.949	-0.340	2.027	-0.303
20	2.125	0.001	2.136	0.000
21	1.859	0.292	1.967	0.335
22	1.284	0.407	1.329	0.325
23	2.486	-0.289	2.636	-0.253
24	2.508	-0.407	2.644	-0.353
25	0.843	-1.298	0.801	-1.456
平均	1.752	-0.204	1.799	-0.170
標準偏差	0.507	0.551	0.535	0.555

注：平均および標準偏差は、項目1～項目25それぞれの推定値を用いて計算した平均および標準偏差である。

出所：2019年調査については筆者作成。2016年調査については、國方(2020)表2に基づき筆者作成。

[表2] 項目パラメータ推定結果 (2019年調査, 等化後)

項目番号	mean & sigma法		mean & mean法	
	項目識別力推定値	項目困難度推定値	項目識別力推定値	項目困難度推定値
1	1.011	-0.202	1.046	-0.200
2	1.244	0.071	1.287	0.063
3	1.636	-0.091	1.693	-0.093
4	0.914	0.177	0.946	0.166
5	1.931	-0.557	1.999	-0.544
6	2.045	-0.742	2.117	-0.723
7	2.250	-1.098	2.328	-1.066
8	1.594	-0.657	1.650	-0.641
9	1.432	0.299	1.482	0.283
10	1.603	-0.128	1.659	-0.129
11	1.747	-0.364	1.808	-0.357
12	2.301	-0.606	2.381	-0.592
13	2.548	-0.832	2.637	-0.810
14	1.603	0.156	1.659	0.145
15	0.935	1.497	0.967	1.441
16	2.356	0.188	2.438	0.176
17	1.686	0.084	1.745	0.076
18	1.674	-0.048	1.733	-0.052
19	1.935	-0.307	2.002	-0.302
20	2.109	0.037	2.183	0.030
21	1.845	0.331	1.909	0.314
22	1.274	0.446	1.319	0.426
23	2.467	-0.255	2.553	-0.252
24	2.489	-0.374	2.576	-0.367
25	0.837	-1.272	0.866	-1.235
平均	1.739	-0.170	1.799	-0.170
標準偏差	0.503	0.555	0.520	0.536

注：表1と同じ。

出所：筆者作成。

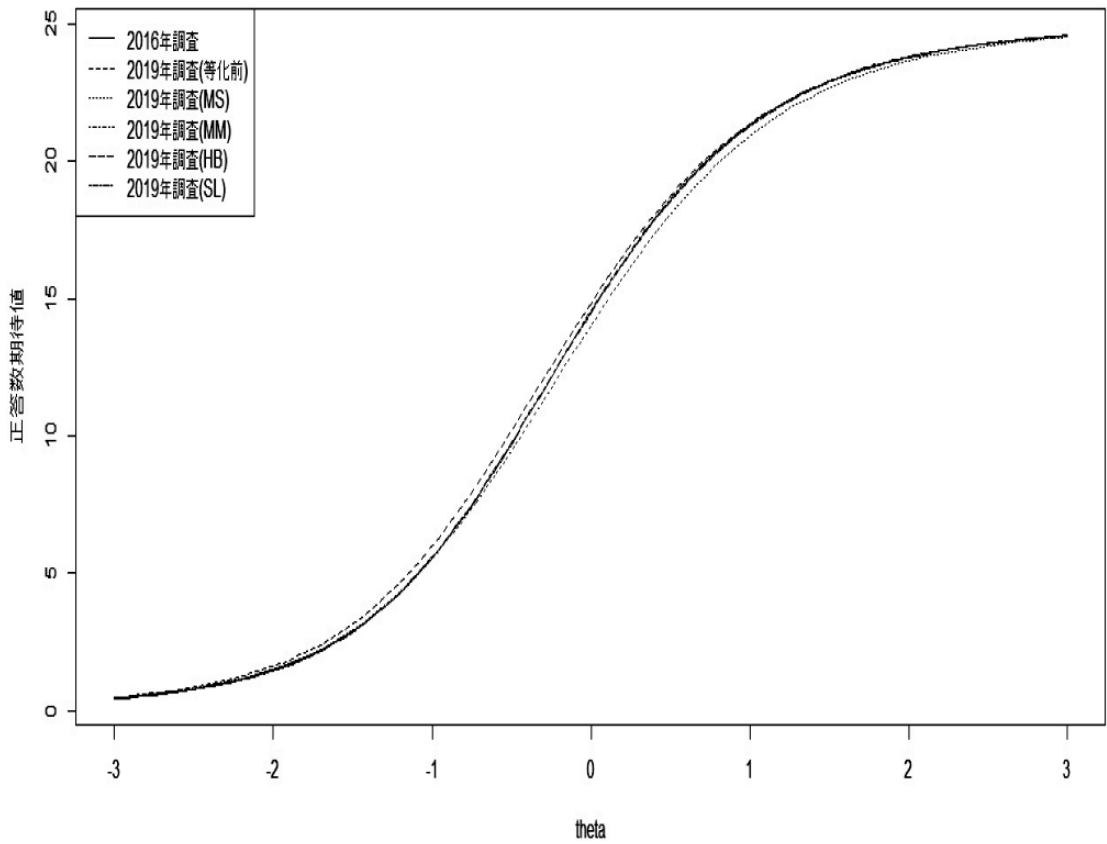
[表2] 続き

項目番号	Haebara法		Stocking & Lord法	
	項目識別力推定値	項目困難度推定値	項目識別力推定値	項目困難度推定値
1	1.037	-0.193	1.043	-0.196
2	1.276	0.072	1.283	0.068
3	1.679	-0.085	1.688	-0.088
4	0.938	0.176	0.943	0.172
5	1.982	-0.540	1.993	-0.540
6	2.099	-0.720	2.110	-0.720
7	2.309	-1.067	2.321	-1.064
8	1.636	-0.637	1.645	-0.637
9	1.470	0.295	1.478	0.290
10	1.645	-0.122	1.654	-0.124
11	1.792	-0.351	1.802	-0.353
12	2.361	-0.588	2.374	-0.588
13	2.614	-0.808	2.629	-0.807
14	1.645	0.155	1.654	0.151
15	0.959	1.462	0.964	1.451
16	2.417	0.187	2.430	0.182
17	1.730	0.085	1.740	0.081
18	1.718	-0.044	1.728	-0.047
19	1.985	-0.296	1.996	-0.298
20	2.164	0.040	2.176	0.036
21	1.893	0.325	1.904	0.320
22	1.308	0.438	1.315	0.432
23	2.531	-0.245	2.545	-0.247
24	2.554	-0.362	2.568	-0.363
25	0.859	-1.236	0.863	-1.233
平均	1.784	-0.162	1.794	-0.165
標準偏差	0.516	0.541	0.519	0.538

注：表1と同じ。

出所：筆者作成。

[図1] テスト特性曲線



注：凡例のMS, MM, HB, SLは, 等化の方法を指す。MSはmean & sigma法を, MMはmean & mean法を, HBはHaebara法を, SLはStocking & Lord法を, それぞれ指す。

出所：筆者作成。