

間主観的自己評価における自動推定を組み入れた適応型自己評価手法の提案

Proposal of an Adaptive self-Assessment Incorporating Automatic Estimation for Inter-Subjective Assessments

眞坂 美江子^{*1}, 渡辺 博芳^{*2}, 宮崎 誠^{*1}

Mieko MASAKA^{*1}, Hiroyoshi WATANABE^{*2}, Makoto MIYAZAKI^{*1}

^{*1} 帝京大学, ^{*2} 電気通信大学

^{*1}Teikyo University, ^{*2}University of Electro-Communications

Email: masaka@ics.teikyo-u.ac.jp

あらまし: ルーブリックやチェックリストを用いた間主観的評価において、継続的な自己評価は学習者に持続的な成長を促す効果が期待されるが、汎用的能力等においては、学習者の各能力をさまざまな観点から評価する必要があり、評価項目が多くなるという課題がある。本研究は自己評価者の回答パターンに応じて最適な評価項目を提示することにより、より少ない評価項目の数で、すべての評価項目を用いた場合と同等の自己評価結果を得る手法を検討した。順序ロジスティックモデルにより、間主観的評価の自己評価結果を他の自己評価結果から予測可能であることを示し、自動推定を組み合わせることで、削減可能な評価項目の数をシミュレーションした結果、8つの評価項目に対し、1~2つを削減できる可能性があることを明らかとした。

キーワード: 適応的自己評価, 間主観的評価, 汎用的能力, 多変量解析

1. はじめに

高等教育においては、従来、知識・技能の修得が重視されていたが、近年はそれらに加えて修得した知識を活用した問題解決力、問題解決のための思考力・判断力・表現力等、特定の分野や職業に限定されず汎用的に活用出来る基礎的な能力（以下、汎用的能力と記す）の修得も重視されている。

知識や技能は、定量的に評価しやすいのに対して、このような能力は、定量的な評価が難しいとされている。ルーブリックやチェックリストを用いた間主観的評価は、客観テストのみでは評価がしづらい定性的な観点の評価をすることが出来る学習達成度の評価手法として着目されている。しかし、汎用的能力を間主観的に評価する際には、学習者の各能力をさまざまな観点から評価する必要があり、評価項目が多くなるという課題がある。

そこで本研究は、汎用的能力のルーブリック等を用いた間主観的評価の自己評価を対象に、自己評価結果の相関関係に着目し、自己評価者の回答パターンに応じて最適な評価項目を提示することにより、より少ない評価項目の数ですべての評価項目を用いた場合と同等の自己評価結果を得る適応型自己評価手法⁽¹⁾を提案する。

2. 汎用的能力自己評価の概要

帝京大学情報電子工学科では、間主観的評価により、学生が自らの汎用的能力を自己評価する取り組みを実施している⁽²⁾。当学科では、9つの能力を観点とする汎用的能力ルーブリックを定義している。学習者はまずチェックリストを用いて各能力の詳細な点を0~2の範囲で評価し、その後、チェックリス

トの結果を参考にしてルーブリックを用い、各能力の総合評価を0~4の範囲で評価する。本研究は、評価項目が多いチェックリストに着目してその削減手法を検討する。

3. 順序ロジスティック回帰分析

本研究は、2020年度から2023年度に4年生のワークショップで実施した9つの汎用的能力の中から、「主体的・継続的な学習力」のデータを用いて順序ロジスティック回帰分析をおこなった。サンプル数は、171件である。分析結果を表1に示す。表1の横軸方向が目的変数である。あるチェック項目の自己評価結果を目的変数、それ以外のチェック項目の自己評価結果を説明変数として順序ロジスティック回帰分析を行い各説明変数の偏回帰係数を記載した。「主体的・継続的な学習力」には、8つのチェック項目が定められている。すべてのチェック項目において、回帰式の精度を示すMcFaddenは、0.16~0.31の範囲であった。McFaddenは、0~0.4が適したモデルであることから、作成したモデルは予測に適したモデルと判断できる。また、モデル全体の χ^2 は、54.27~103.04、p値はすべて、0.001以下であり、回帰式は有意であると言える。

表中の網掛けは、偏回帰係数が、1%もしくは5%有意となったものである。目的変数によって、1~5個程度の偏回帰係数が有意な係数となった。さらに、表2に、正解率とF値をまとめた。F値の計算に必要なサンプル数が得られなかったセルは、「—」と記載している。正解率は0.62~0.96と全体的に良い正解率となっているが、F値に大ききばらつきがみられる。F値は、一般的に0.7以上が良いモデルとされ

表1 偏回帰係数

		n=171							
目的変数		ch1	ch2	ch3	ch4	ch5	ch6	ch7	ch8
McFadden		0.29	0.30	0.28	0.27	0.26	0.31	0.21	0.16
χ^2		86.21	93.71	98.67	81.06	90.05	103.04	58.44	54.27
説明変数	ch1	—	0.79	0.36	0.89	0.23	0.75	1.30	-0.54
	ch2	0.78	—	0.74	0.79	0.05	0.79	0.10	0.65
	ch3	0.41	0.70	—	0.02	1.24	0.38	0.18	0.87
	ch4	0.86	0.87	0.11	—	0.79	0.48	0.52	0.39
	ch5	0.12	0.06	1.22	0.78	—	1.06	-0.18	-0.33
	ch6	0.71	0.79	0.44	0.40	1.16	—	0.42	0.57
	ch7	1.34	0.06	0.14	0.54	-0.23	0.40	—	0.63
	ch8	-0.42	0.72	0.82	0.35	-0.26	0.43	0.56	—
			p<0.05		P<0.01				

表2 正解率およびF値

		正解率			F値		
目的変数		0:できていない	1:ある程度できていない	2:できている	0:できていない	1:ある程度できていない	2:できている
		ch1	0.94	0.71	0.75	0.17	0.67
ch2	0.94	0.71	0.75	0.44	0.69	0.74	
ch3	0.82	0.62	0.77	0.48	0.60	0.68	
ch4	0.95	0.67	0.72	0.18	0.68	0.69	
ch5	0.85	0.64	0.77	0.42	0.64	0.70	
ch6	0.92	0.69	0.77	0.59	0.69	0.72	
ch7	0.96	0.71	0.74	—	0.65	0.77	
ch8	0.89	0.68	0.75	0.34	0.69	0.69	

ていることから、F値0.7以上に着目すると、自己評価「2:できている」においては、半数以上のF値が0.7以上となっており、正解率も7割を超えている。一方で、「0:できていない」については、正解率は高いが、F値が低いものが多数見られる。

4. 適応型自己評価手法の提案

以上結果から、順序ロジスティック分析による推定精度には、ばらつきがあるため、適切な推定が可能な回答パターンに限定して未評価の自己評価結果を自動推定すべきである。また、間主観的自己評価は、評価者が評価項目を見ることにも意義があるため、特定の評価項目のみを毎回提示することは適切でない。評価者が自己評価を繰り返し行う際には、評価項目がまんべんなく提示されることが望ましい。そこで図1の手順によるチェック項目の自己評価数削減手法を提案する。システムが評価項目を毎回ランダムに提示しながら、未評価の評価項目の自己評価結果を自動推定することで、学習者が実際に評価すべき項目数を削減する手法である。

5. チェック項目削減シミュレーション

図1の提案手順に従いシミュレーションを行い、どれくらいのチェック項目が削減可能であるか評価した。2022~2023年度の自己評価データを用いて順序ロジスティック回帰モデルを作成し、2024年度の自己評価データを用いてシミュレーションした。シミュレーションデータ数は45件である。

すべてのチェック項目が自己評価済、もしくは、自動推定済となるまで、①~②の手順を繰り返す。
①自己評価済、および自動推定済の評価を除くチェック項目の中からランダムにチェック項目を選択し、自己評価者に評価させる。
②表1を参照し、有意な偏回帰係数を持つチェック項目の自己評価が全て完了し、かつ、F値が0.7以上となるチェック項目があれば、その自己評価結果を自動推定する。

図1 適応型自己評価手法

表3 適応型自己評価シミュレーション結果

自己評価項目数	自動推定項目数	人数
6	2	17
7	1	87
8	0	121
合計		225

チェック項目の提示はランダムに行うため、同じ自己評価データに対して5回シミュレーションを行った。表3に自動推定により削減できたチェック項目の数ごとにまとめた件数を示す。すべてのチェック項目について自己評価が必要となったものは121件であり、約半数のケースで、学習者が自己評価するチェック項目の数を1~2つ削減することができた。すべて自己評価を行った結果と、自動推定を組み合わせて評価を行った結果の一致率は、99.4%であり、自動推定を一部導入しても、自己評価結果精度の大幅な低下は見られなかった。

6. まとめ

本研究は、間主観的自己評価における自己評価項目数の削減を目指し、多変量解析による自己評価結果の自動推定手法を提案した。シミュレーションの結果、約半数のケースにおいて、数個ではあるが、学習者が自己評価すべき項目を削減し、すべての評価項目を用いた場合と同等の自己評価結果を得ることができた。今回の分析は、帝京大学情報電子工学科が掲げる汎用的能力の一部のみであり、今後、提案手法が他の能力に対しても適応可能であるか、さらなる分析が必要である。

謝辞

本研究はJSPS 科研費24K06217、および帝京大学先端総研インキュベーション助成金23-98の助成を受けている。

参考文献

- (1)真坂美江子, 宮崎誠, 渡辺博芳: 間主観的な評価における適応型自己評価に関する研究のグランドデザイン, 情報教育シンポジウム論文集, pp.328-335 (2024).
- (2)渡辺博芳, 荒井正之, 他: 汎用的能力評価のためのルーブリックとチェックリストの提案, 情報教育シンポジウム論文集, pp.30-37 (2019).