

情報環境利用に関する満足度データの項目反応理論による検討

岩間 徳兼¹ 木村 好美² 石田 崇¹ 須子 統太¹ 末松 大¹
 早稲田大学メディアネットワークセンター¹ 早稲田大学文学学術院²
 n.iwama0119@aoni.waseda.jp

概要：2011 年度に実施された早稲田大学情報利用アンケートより得られた情報環境利用に関する満足度を尋ねた項目への回答データに、項目反応理論に基づくモデルを適用し、文理やキャンパスによる各項目の特徴の違いについて検討した。また、属性ごとの結果を統合して布置図として表し、満足度に関連する項目を見出す方法を示した。

1 情報環境利用アンケート

早稲田大学（以下、本学）メディアネットワークセンター（以下、MNC）では、コンピュータールームの設備、IT ツールなどの情報環境利用に関するアンケート調査を教員、職員、学生を対象として年に一度実施し、その結果を情報環境の改善のために活用している。このアンケートは情報環境利用アンケートと呼ばれ、集計結果は本学の Web ページ上で公開されている [1]。

アンケートは IT ツールの使用方法、MNC 設置科目・提供サービスに対する意見、情報環境利用における満足度を尋ねる項目などから構成されている。ここでは、情報環境利用に対する満足度を問う項目への学生（学部生および大学院生）の回答データを取り上げ、情報環境への満足度の測定という計量心理学的観点からそれらの項目や得られる結果について検討を行う。

2 問題と目的

情報環境利用アンケートの調査目的の一つは、情報環境利用に対する満足度を把握することである。これまでは主に、CS（Consumer Satisfaction, 顧客満足）ポートフォリオ分析¹を利用して、その検討がなされてきた。その際、学生を何らかの属性で区別して検討することは行われてこなかった。観測対象を層化した視点から検討することは、全体をひとまとめにした場合に見逃してしまう特徴を捉える際に有効な

¹CS ポートフォリオ分析はマーケティングリサーチにおいてしばしば利用される手法である。基本的な考え方は、個々の事柄（端末室 PC のスペック、科目登録のシステムなど）に対する満足度とそれらの総体（情報環境全体）に対する満足度を同時に尋ね、総体に対する個々の事柄の重要度と個々の事柄への満足の大きさの二次元上に個々の事柄を布置することで、各事柄の傾向を捉えようというものである。

手段である。情報環境に対する理解やツールに触れる機会は文系学生と理系学生で異なるであろうし、コンピュータールームなどの施設や LAN の設備はキャンパスごとに異なるため、この調査データの検討にあたっては、情報環境に対する全学での満足度だけでなく層別の満足度も捉えることが望ましい。

しかし、CS ポートフォリオ分析においては、重要度と満足度はそれぞれ偏差値によって数値化されているため、各層内での検討は可能であっても層間で検討を行ったり、それらを統合して検討を行ったりすることは不可能である。そこで、ここでは満足度の測定という観点から計量心理学的手法である項目反応理論（Item Response Theory, 以下 IRT）[2] を利用し、学生の情報環境利用に対する満足度を層ごと、全体の二つの視点から考察することを試みる。

3 項目反応理論の概要

IRT は尺度（テスト）の信頼性や妥当性の定式化などで利用されている古典的テスト理論と対比させ、現代テスト理論と呼ばれるものである。IRT では、満足度などの測定対象を観測不可能な潜在変数として捉え、各回答者の特性値（潜在変数の値）と各項目の項目母数（項目特徴を示す母数）によって項目に対する反応確率が規定されるというモデル化を行う。また、そのようなモデル化に基づき、特性値と項目母数のどちらに関しても、異なった機会や場所で回答された結果を比較することが可能である。その優れた性質から、TOEIC や日本語能力試験、IT パスポート試験等、多くのテストで IRT による運用が行われている。回答の形式（二値・多値）、組み込む母数の数、利用

する確率分布などによって様々なモデルが提案されているが、ここでは一般的な、二値反応に対する二母数ロジスティックモデルを利用することとする。

回答者 i の項目 j に対する反応を u_{ij} と表すことにする。 u_{ij} は当該項目に正答（もしくは肯定的反応）ならば1、誤答（もしくは否定的反応）ならば0をとるものとする。この時、その項目に対する尺度値 θ_i に基づく確率分布を、二母数ロジスティックモデルを利用して

$$P_i(u_{ij} | \theta_i, a_j, b_j) = \frac{\exp\{-Da_j(\theta_i - b_j)(1 - u_{ij})\}}{1 + \exp\{-Da_j(\theta_i - b_j)\}} \quad (1)$$

と表現する。ここで、 a_j, b_j はそれぞれ項目 j の識別力、困難度を表す母数である。識別力が大きい項目は特性値の違いが正答率に反映されやすい項目であり、困難度が大きい項目は正答率が低くなりやすい項目であるとおおよそ理解できる。 D は尺度因子であり、ロジスティック分布を正規分布へ近似させるため1.7とされることが多い。また、尺度値は $\theta_i (\theta_i \sim N(0, 1))$ であると仮定する。なお、(1)式において $a_j = a$ としたモデルは一母数ロジスティックモデルとして知られ、識別力については項目間で共通であるモデルを表す。

4 特異項目機能

回答者の属性の違いによる項目への反応の違いを捉えるために、特異項目機能 (Differential Item Functioning, 以下 DIF) の観点を採用する。IRT における DIF は異なる群に所属し、同じ能力レベルを有するにもかかわらず当該項目に対する正答確率が異なる状態のことを指す。文化間の比較を目的として DIF の検出を利用することも多い [3]。特性値の測定という側面から見ると、DIF 項目の存在は属性という局外母数が特性値測定に影響を及ぼすことを意味し、それがバイアスのかかった測定につながりうるという点で望ましくないことである。したがって、測定値自体に興味がある場合には、DIF 項目を除いて特性値の推定を行ったほうがよい。

IRT を利用した DIF の分析手法としては、Raju 法 [4, 5] や Lord 法 [2] などがあり、これらは二つの群の比較に利用可能な方法である。本稿においては、Lord 法を多群比較へと拡張した一般化 Lord 法 [6] を利用する。この方法

では

$$Q_j = (Cv_j)'(C\Sigma_j C')^{-1}(Cv_j) \quad (2)$$

のように検定統計量の値を構成する。ここで、 v_j は項目 j に関して全ての群の等化²後の項目母数推定値をつなぎ合わせたベクトルである。 Σ_j はブロック対角行列であり、その対角要素の行列は各群の項目母数推定量の共分散行列の推定値である。また、 C は様々な形の検定を構成するための計画行列である。項目母数推定量の漸近正規性を利用すると、 Q_j は C の階数と同じ自由度の χ^2 分布に従う確率変数の実現値とみなせる。その χ^2 分布上での Q_j に対応する有意確率 (p 値) を利用することで、項目 j が DIF 項目かそうとは言えないかという結論を導くことができる。

5 分析結果

以降では、2011 年度に実施された情報環境利用アンケートから得られた情報環境利用に対する満足度を問う項目への学生の回答データを IRT を用いて分析した結果を示す。

5.1 アンケート調査 (学生対象) 概要

学生を対象とした調査の概要は以下に示す通りである。

実施期間：2011 年 11 月 15 日から 11 月 28 日

実施方法：本学が提供する Waseda-net ポータルのアンケート機能を使用し、メールにて学生に通知し、Web 経由でアンケートへの回答を求めた。

対象者数：61,663 名

回答者数：3,422 名

情報環境利用に対する満足度を問う項目は表 1 の通りである。回答は、「そう思わない」、「ややそう思わない」、「どちらでもない」、「ややそう思う」、「そう思う」、「知らない・使ったことがない」、から一つを選択する多肢選択形式であった。

5.2 項目分析

表 1 の最右列には各項目における認知率 (全回答者のうち「知らない・使ったことがない」以外の回答をした者の比率) を掲載した。まず、問 1.7 は情報環境全体についての満足度であり、

²等化は異なる機会や場所で得られたテストの結果を比較可能にするための手続きのことである。

表 1 情報環境利用に対する満足度を問う 24 項目

項目	設問内容	認知率
問 1.7	早稲田大学が提供する情報環境に満足していますか。	0.99
問 2.1	学内共通端末室環境に満足していますか。	0.96
問 2.4	学内共通端末室の開室時間は適当だと思いますか。	0.94
問 2.5	端末室の PC について、スペックは十分だと思いますか。	0.94
問 2.8	端末室の PC 設置台数は十分だと思いますか。	0.95
問 2.9	端末室に設置されたプリンタの台数は十分だと思いますか。	0.91
問 2.12	学内の有線ネットワークについて回線速度は十分だと思いますか。	0.67
問 2.13	学内の情報コンセントの設置数について十分だと思いますか。	0.73
問 2.14	学内の無線ネットワークについて回線速度は十分だと思いますか。	0.83
問 2.15	学内の無線 LAN アクセスポイントの設置数について十分だと思いますか。	0.70
問 2.16	端末室以外の教室の AV 設備について十分に整備されていると思いますか。	0.85
問 3.1	Waseda-net ポータルに満足していますか。	1.00
問 3.13	Waseda-net メールに満足していますか。	1.00
問 3.17	Course N@vi に満足していますか。	0.98
問 3.27	文献管理用ソフトウェアまたは文献管理サービス (Refworks 等) に満足していますか。	0.41
問 3.28	早稲田大学リポジトリ「DSpace」に満足していますか。	0.22
問 3.29	Web 科目登録システムに満足していますか。	0.97
問 3.34	Web 科目登録システムについて、レスポンス (応答速度) について十分だと思いますか。	0.98
問 4.2	MNC 設置科目全般 (内容、レベル、ラインナップ) に満足していますか。	0.31
問 5.1	ネットワーク型授業 (オンデマンド授業、CCDL 授業) 全般に満足していますか。	0.66
問 5.2	オンデマンド授業に満足していますか。	0.66
問 5.8	CCDL 授業に満足していますか。	0.15
問 6.2	「QuonNet」に満足していますか。	0.09
問 6.6	「緊急お知らせサイト」に満足していますか。	0.56

IRT による分析では必要ないため、分析対象の項目から除外した。また、認知率の低い項目は満足度の測定にとって負の影響を与え、IRT によるモデル推定時の標本サイズの減少にもつながるので分析対象から除外した (9 項目: 問 2.12, 問 3.27, 問 3.28, 問 4.2, 問 5.1, 問 5.2, 問 5.8, 問 6.2, 問 6.6)。そして、「そう思わない」、「ややそう思わない」、「どちらでもない」の三つを 0, 「ややそう思う」、「そう思う」の二つを 1, 「知らない・使ったことがない」を欠損扱いとしてデータの加工を行った。その結果、回答者数 909 名、項目数 14 個から分析対象となるデータセットが構成された。得られたデータについて、14 項目で形成されるテストの信頼性³を算出したところ、KR-20⁴は 0.84 となり、信頼性の高さが示唆された。

5.3 層化分析-DIF 検出および DIF 項目の特徴の把握-

加工されたデータについて項目反応モデルを当てはめ、さらに DIF の検出を行った。分析には、統計ソフト R[7] のパッケージ difR に含

³測定の信頼性とも呼ばれ、測定値が測定時に偶然的作用する要素によって受ける影響の弱さを意味する。

⁴内的整合性 (テストを構成する項目の同質性) からの信頼性の推定値であり、1 に近いほど信頼性が高いと判断される。

まれる関数を利用した。

表 2 文理属性による項目母数推定値および DIF 検出結果

項目	文系		理系		DIF 検出	
	識別力	困難度	識別力	困難度	χ^2 値	p 値
問 2.1	1.422	0.077	1.699	-0.151	1.543	0.462
問 2.4	0.866	-0.770	1.036	-0.711	0.623	0.732
問 2.5	0.924	0.445	1.017	0.233	0.995	0.608
問 2.8	0.995	0.098	0.713	0.362	3.487	0.175
問 2.9	0.768	1.228	0.802	-0.309	47.643	0.000
問 2.13	1.185	1.018	1.788	0.872	3.413	0.182
問 2.14	0.798	0.291	1.149	0.024	3.996	0.136
問 2.15	1.240	1.061	1.383	1.330	1.712	0.425
問 2.16	0.910	-0.301	0.985	-0.035	1.389	0.499
問 3.1	1.244	-0.283	1.456	-0.346	0.733	0.693
問 3.13	1.053	-0.326	1.438	-0.275	2.102	0.350
問 3.17	1.360	-0.044	1.835	-0.124	2.175	0.337
問 3.29	0.905	-0.625	1.579	-0.682	7.029	0.030
問 3.34	0.979	-0.751	1.266	-0.663	1.385	0.500

文理属性による層化

まず、回答者を所属学部により、文系学部と理系学部に分け (文系 588 名, 理系 321 名), 文理属性の観点から DIF の検出を行った (表 2)。その結果、問 2.9 と問 3.29 については χ^2 検定の結果がそれぞれ 0.1%, 5% の有意水準で有意となり、DIF 項目であると認められた。これら二つの項目について等化後の項目反応関

数⁵ (Item Response Function, 以下 IRF) を図1と図2に示す(破線が文系, 実線が理系)。

問2.9は「端末室に設置されたプリンタの台数は十分だと思いますか。」という内容の項目である。文系における項目母数の推定値は識別力0.768, 困難度1.228であり, 理系における推定値は, 識別力0.802, 困難度-0.309であった。文理の層化において識別力には大きな違いはないものの困難度の違いが顕著である。このことから, 文系の方が理系よりも困難度が高い, すなわち, 文系学生のほうがプリンタの台数が十分であるとは答えにくいことが分かる。

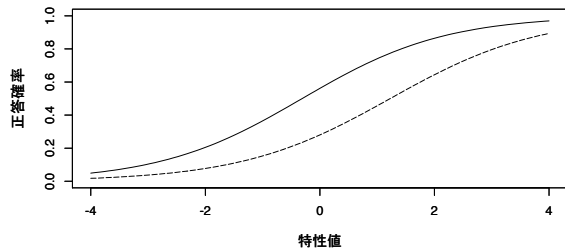


図1 問2.9のIRF(文理属性)

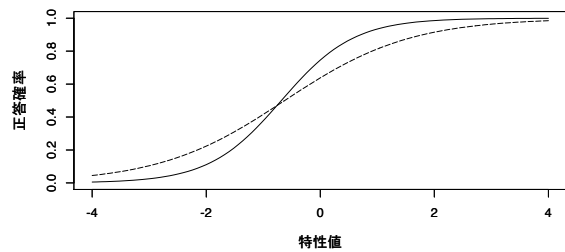


図2 問3.29のIRF(文理属性)

続いて, 問3.29は「Web科目登録システムに満足していますか。」という内容の項目である。文系における項目母数の推定値は識別力0.905, 困難度-0.625であり, 理系における推定値は識別力1.579, 困難度-0.682であった。この項目については, 識別力における文理間での違いが明確であり, そのことによって文理でのIRFが交差している(図2)。理系のIRFのほうが立ち上がり急であるので, 満足度の違いが肯定的な回答の生起確率に大きく反映される, すなわち, 理系学生は文系学生よりもWeb科目登録システムへの満足度において敏感であ

⁵ある項目における特性値と正答確率の対応を示す関数であり, 横軸に特性値, 縦軸に正答確率をとった座標上において, 連続で単調増加の曲線を示す。

ることがわかる。理系学生はインターネット等に触れる機会が比較的多く, 様々なシステムの利用経験があり, それが敏感さにつながっているのかもしれない。

キャンパス属性による層化

次に, キャンパスの違いから回答者を早稲田, 戸山, 西早稲田, 所沢, 北九州, 本庄のカテゴリへと分類した(早稲田334名, 戸山141名, 西早稲田321名, 所沢73名, 北九州37名, 本庄3名)。北九州と本庄の属性を有する回答者は少なく, キャンパス属性による層化の分析からは除外することとし, 早稲田, 戸山, 西早稲田, 所沢の四つのキャンパスに関してDIFの検出を行った(標本サイズは869)。所沢キャンパスの回答者数が少ないこともあり, 項目反応モデルとして一母数ロジスティックモデルを採用した。母数の推定値およびDIF検出の結果を表3に示した。なお, 識別力母数の値は1に固定されており, 表には示していない。

表3 キャンパス属性による困難度推定値およびDIF検出

項目	キャンパス				DIF検出	
	早稲田	戸山	西早稲田	所沢	χ^2 値	p値
問2.1	0.037	0.066	-0.139	0.294	2.181	0.536
問2.4	-0.538	-0.932	-0.669	-0.593	2.306	0.511
問2.5	0.426	0.247	0.325	0.444	0.675	0.879
問2.8	0.161	-0.151	0.372	0.519	5.554	0.135
問2.9	0.835	1.267	-0.186	2.054	60.245	0.000
問2.13	1.043	0.985	1.355	1.077	3.045	0.385
問2.14	0.192	0.427	0.108	0.294	1.734	0.630
問2.15	1.079	1.267	1.773	1.163	10.789	0.013
問2.16	-0.256	-0.262	0.047	-0.368	3.508	0.320
問3.1	-0.333	-0.115	-0.375	-0.746	3.013	0.390
問3.13	-0.412	-0.225	-0.280	-0.147	1.064	0.786
問3.17	0.037	-0.151	-0.108	-0.221	1.111	0.775
問3.29	-0.666	-0.225	-0.858	-0.592	6.215	0.102
問3.34	-0.634	-0.605	-0.720	-0.903	0.838	0.840

キャンパスの観点からは, 問2.9および問2.15がそれぞれ有意水準0.1%, 5%でDIF項目として検出された。図3に問2.9「端末室に設置されたプリンタの台数は十分だと思いますか。」のIRFを示した。左から西早稲田(困難度-0.186), 早稲田(同0.835), 戸山(同1.267), 所沢(同2.054)となっており, 一番推定値の低い西早稲田と高い所沢は他の二つのキャンパスとの差が大きいことがわかる。理工系の学生の多い西早稲田キャンパスについては, 項目に対して肯定的な回答をしやすいが, 所沢キャンパスの学生は肯定的な回答を示しに

くということが認められる。続いて、図4は問2.15「学内の無線LANアクセスポイントの設置数について十分だと思いますか。」のIRFである。左から、早稲田(困難度1.079)、所沢(同1.163)、戸山(同1.267)、西早稲田(同1.773)のIRFを表している。この項目に関しては、西早稲田のIRFが他の三つから離れており、西早稲田の学生からは、他のキャンパスの学生よりも無線LANアクセスポイントの設置数について満足しているという回答が得られにくいことが読み取れる。

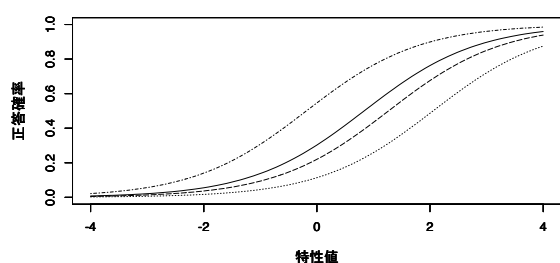


図3 問2.9のIRF(キャンパス属性)

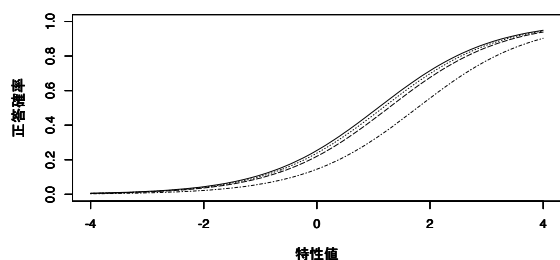


図4 問2.15のIRF(キャンパス属性)

5.4 IRTによる統合的CSポートフォリオ分析

困難度は各項目の満足度(の低さ)、識別力は各項目における潜在的な満足度と満足率との関連の強さを意味しているから、困難度や識別力を利用してIRTの文脈からCSポートフォリオ分析を行うことが可能であると考えられる。ここでは、二母数モデルを採用した文理属性での結果を利用して、CSポートフォリオ分析を行う。まず、文系と理系において得られた等化済みの項目母数(表2)に関して、DIF項目であるとは結論付けられなかった項目についてはその平均値を統合された項目母数の値とし、DIF項目であると判断された項目については文系、理系をそれぞれ、別の推定値として扱うこ

ととした。そして、識別力と困難度それぞれで標準化を施した(表4)。その結果4に基づいて、識別力を横軸、困難度を縦軸とした座標上に各項目をプロットしたものが図5である。

図中において右上のエリアには全体として統合された二つの項目(問2.13, 問2.15)が布置されている。右上にある項目は満足度との関わりが強いにも関わらず、満足だという反応が得られにくかった項目である。以上より、この項目の内容である「情報コンソールの設置数」と「無線LANアクセスポイントの設置数」については、全学で改善する余地があるという示唆が得られる。また、右下のエリアには理系に関する一つの項目(問3.29(理))が布置されている。右下に位置する項目は満足度との関わりが強く、満足だという反応が得られやすかった項目である。したがって、この項目の内容である「Web科目登録システム」は、理系の学生において満足度の高さに寄与する要因となっていると判断することができる。

表4 文理属性に関して統合した項目母数(標準得点)

項目	識別力	困難度
問2.1	1.390	-0.070
問2.4	-0.681	-1.160
問2.5	-0.613	0.513
問2.8	-1.011	0.344
問2.13	1.137	1.452
問2.14	-0.604	0.232
問2.15	0.543	1.839
問2.16	-0.692	-0.273
問3.1	0.674	-0.499
問3.13	0.318	-0.478
問3.17	1.515	-0.142
問3.34	-0.098	-1.108
問2.9(文)	-1.304	1.891
問2.9(理)	-1.188	-0.492
問3.29(文)	-0.837	-0.980
問3.29(理)	1.451	-1.069

6 まとめ

本稿では、本学において実施されているアンケートより得られた情報環境利用に対する満足度データをIRTを用いて分析し、結果の検討を行った。DIFの検出法を用いることで、これまでの分析では行われなかった層別の分析が可能となり、どの項目について層間で違いが存在するのかを知ることができる。また、それらの手法によって、これまでCSポートフォリオ分析では不可能であった層間での結果の比較や統

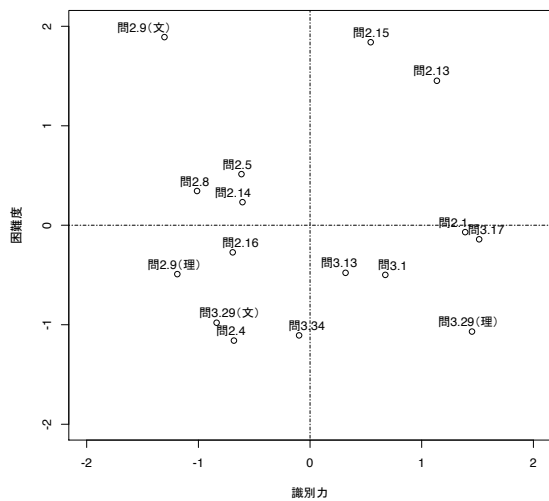


図5 文理属性による各項目のプロット

合が可能であることが示された。今回の分析例においては回答者属性による層化を行ったが、年度間の比較においても本手法を適用することができる。情報環境利用に対する満足度を改善すべく何らかの対策を講じた場合には、その効果を年度間の比較によって把握することに意味がある。これまでのCSポートフォリオ分析では年度間の比較も無理であったが、IRTとDIF検出法を組み合わせることで年度に関して普遍的な部分と特異な部分を同時に検討することが可能となる。

従来のCSポートフォリオ分析との比較や層間での結果の統合の方法についての検討は今後の課題としたい。

参考文献

- [1] 情報環境利用アンケート集計結果(早稲田大学教務部情報企画課)、<http://www.waseda.jp/wits/data/questionnaire/index.html>、2012/11/09/15:45 アクセス
- [2] Lord, F. M., 「Applications of item response theory to practical testing problems」、Lawrence Erlbaum Associates, 1980
- [3] 田崎勝也, 「社会科学のための文化比較の方法 等価性とDIF分析」、ナカニシヤ出版、2008
- [4] Raju, N. S., 「The area between two item characteristic curves」、Psychometrika 53, 495-502, 1988

- [5] Raju, N. S., 「Determining the significance of estimated signed and unsigned areas between two item response functions」、Applied Psychological Measurement 14, 197-207, 1990
- [6] Kim, S.-H., Cohen, A.S., and Park, T.-H., 「Detection of differential item functioning in multiple groups」、Journal of Educational Measurement 32, 261-276, 1995
- [7] Ihaka, R., and Gentleman, R., 「A language for data analysis and graphics」、Journal of Computational and Graphical Statistics 5, 299-314, 1996