
趣味・音楽の好み・ネットメディア利用に 基づいた若年女性文化の分類

—潜在クラスモデルとベータ回帰モデルを用いて—

木 田 勇 輔

要約

人が類似性に基づいて集団形成を行うことはよく知られているが、現代社会において類似性を判断する基準の一つは趣味や関心であろう。本稿では①趣味・音楽の好み・ネットメディア利用の回答から若年女性においてどのような文化的分類が可能であるのか、②上記の文化的分類はどのような社会属性的要因と関連しているのかという2点について、潜在クラスモデルとベータ回帰モデルを用いた分析を行う。分析にあたっては、MCMCを用いたベイズ推定を利用する。データは20代の女性を対象としたウェブ調査から得たものである。弱事前分布を用いた潜在クラス分析の結果からは6つのクラスが析出され、それぞれに対して一般趣味層（35.5%）、少趣味層（25.1%）、アクティブ＝オタク層（13.0%）、オタク層（10.7%）、アクティブ層（10.5%）、雑食層（5.3%）という解釈を行った。また、回帰係数と切片に弱情報事前分布を用いたベータ回帰分析の結果からは、一般趣味層では年齢（+）と非正規雇用（-）、アクティブ＝オタク層では年齢（+）、オタク層では結婚（-）と非正規雇用（+）と雇用その他（+）、雑食層では正規雇用（+）がそれぞれのクラス所属と関連する可能性が高いことが分かった。本稿の結果からは先行研究で繰り返し言及されてきた社会階層と文化的分類の関連性は確認できなかった。ライフコースや雇用形態など、余暇活動に大きな影響を与える可能性のある要因にも注目する必要がある。ただし、本研究は試論としての性格が強く、データの収集・処理・分析の方法には改善の余地がある。

1. 問題の所在

人が類似性（homophily）に基づいて集団形成を行うことはよく知られているが、現代社会において類似性を判断する基準の一つは趣味や関心であろう。現代社会において趣味や関心を同じくする人々はインターネットで情報交換を行い、ときに現実空間でも行動をともにする。一部の社会学者はこれを趣味縁と呼んで研究の対象としてきた（浅野 2011; 辻 2015）。社会学の伝統的な用語を

用いるとすれば、こうして作られた集団はスタイルや価値観などを含んだサブカルチャー（下位文化）を共有するであろう（Hebdige 1979=1986; Fine 1983; Jenkins [1992]2013; Jenkins 2006=2021; 難波 2007; Haenfler 2014; Haenfler 2016）。社会学の伝統においてが、サブカルチャーとは外部の社会から区別される部分的な集団の価値観・規範、アイデンティティ、行動様式といった文化的な要素を指すことが多いが¹⁾、現代の若者においてそのような集団はもっぱら趣味や関心に基づいて形成されているように思われる。

とくに、近年の日本社会においては趣味や関心

に基づいた活動を活発に行う若年女性の存在感が際立っている。2010年代以降、とくにメディア研究の領域において若年女性の趣味や関心をテーマとした論文集やテキストが続々と出版されている（馬場・池田編 2012; 吉光・池田・西原編 2017; 田中編 2021）²⁾。個別的なテーマについて見れば、ファンダム研究の領域で女性を主たる対象とした重要な研究業績も出版されている（東 2015; 須川 2021）。若年女性の趣味や関心に関する研究は、近年ますます活況を呈していると言えよう。

その一方で、これまで挙げた研究のほとんどが定性的な調査や研究に基づいたものであり、ソーシャルメディアの利用に着目した若年女性の類型化を行った三浦・天笠（2019）のような例外を除けば、若年女性の趣味や関心に基づく文化的分類の定量的な研究はこれまで十分に行われてこなかった。その三浦・天笠（2019）においても、提示されている類型化がどのような基準で行なわれたものであるのかは計量的な視点からは十分に明らかではない。

本稿では趣味に関する社会学的研究、とくに計量社会学的な研究に着目して、若年女性の趣味・関心に基づく文化的分類について明らかにしたい。研究手法面での本稿の大きな特徴は、MCMC（マルコフ連鎖モンテカルロ法）を用いたベイズ統計モデリングを活用しつつ、潜在クラス分析と回帰分析を組み合わせる点にある（Chan and Goldthorpe 2005; 2007; 藤原ほか 2012; White and Murphy 2014）。そして、この手法により①趣味・関心の回答から若年女性においてどのような文化的分類が可能であるのか、②上記の文化的分類はどのような社会属性的要因と関連しているのかという2点について計量的な視点から明らかにしたい。

2. 先行研究の整理と本研究の関心

趣味や関心を共有する集団を研究対象とする際に、社会学では長い歴史を持っている「テイスト」（taste、趣味や嗜好）に関する研究の系譜は極めて重要である。フランスの社会学者であるP. Bourdieuはその著書『ディスタンクシオン』において、個人のテイストが所属する社会階層（階級）と密接に結びついていることを明らかにした（Bourdieu 1979=1990）。その一方で、アメリカ合衆国の社会学者であるR. A. Petersonは、エリートが高級志向（highbrow）の文化を必ずしも好むわけではなく、むしろエリートは高級志向の文化と庶民志向の文化（lowbrow）の双方を享受する雑食性（オムニボア、omnivore）を持つと論じた（Peterson 1992; Peterson and Kern 1996）。

BourdieuやPetersonに影響を受けた研究は日本でも数多く行われている（近藤 2011; 小藪・山田 2013; 小林・大林 2016; 北田・解体研 2017; 磯・竹ノ下 2018; 片岡 2019）。これらの研究では、定量的調査による人々の趣味や関心の把握が重視されてきた。たとえば、2015年SSM調査においても、文化活動の頻度が質問項目として設定されている。高級文化の指標としては「クラシック音楽の音楽会・コンサートへ行く」、大衆文化の指標としては「パチンコをする」が設定されており、その活動頻度として「週に1回以上」から「ここ数年間したことはない」までの5段階を問うものとなっている（磯・竹ノ下 2018）。また、北田らの研究では「アニメ」「食べ歩き」などの若年層を念頭に置いた文化活動を20種類配置して、趣味としてあてはまるものを多重回答式で問うている（北田・解体研 2017）。

Bourdieuの系譜にある研究では多重対応分析がしばしば用いられる。多重対応分析とは、主成分分析の応用的な手法の一つであるが、共分散行列

ではなくカイ二乗距離を用いて特異値分解を行う点にその特徴がある。多重対応分析では分析の結果得られた軸(イナーシャ)とその負荷量をプロットするが、そのグラフが「社会空間」を反映するものとして解釈がなされることが多い(Bourdieu 1979=1990; 磯・竹ノ下 2018)。その一方で、PetersonとKernは回帰分析を用いて文化的オムニボア仮説を検証しているが、回帰分析の利用はこれに続く研究でも踏襲されている(Peterson and Kern 1996; 小林・大林 2016; 北田・解体研 2017; 片岡 2019)

ここで本稿が注目したいのが、T. W. ChanとJ. H. Goldthorpeによる潜在クラス分析の導入である³⁾。潜在クラス分析とは、測定されたカテゴリカル変数群の背後に潜在クラスが存在すると考え、測定された変数と潜在クラスとの関連性を推定する手法である(McCutcheon 1987)。手法の面で言えば、クラスターリングやパターン認識と呼ばれる手法群の一つと見ることもできる。ChanとGoldthorpeは音楽や舞台芸術・映画など回答者の趣味・関心に潜在クラス分析を適用することで、選択のパターンの背後にある社会集団の分類を試みている(Chan and Goldthorpe 2005; 2007)。この手法は、観測された変数の背後に潜在的な集団が存在すると仮定される場合には非常に有効である。

本稿では潜在クラス分析を用いることで、若年女性(ここでは20歳代を想定する)において趣味・嗜好の方向性を同じくする集団がどのような形で見出せるかを考えてみたい⁴⁾。一般的な感覚として、人々を趣味や嗜好の面で集団に分類することができるということは、よく知られているであろう。難波(2008)が歴史社会学的な研究を行っているように、日本においても数多くの「族」や「系」が生まれてきた。たとえば、アンノン族にルーツを持つような旅行やファッションに関心を持つ若年女性を一つの集団として捉えることもできるであろうし、アニメやマンガといったオタク系のコ

ンテンツに関心を持つ若年女性を一つの集団として捉えることもできるであろう⁵⁾。さらに、三浦・天笠(2019)が注目しているように、こうした集団形成はソーシャルメディアの利用とも関連しあっているであろう。

ただし、このような分類が単にステレオタイプを具象化しただけのものになってしまう可能性もあることには注意しなければならない。現代社会においても若年女性を表象に基づいた類型に安易に当てはめようとする言説は根強く存在している。しかし、こうした類型化はステレオタイプに過ぎないこともあるため、経験的な社会科学の研究ではこうした言説を相対化する視点も求められるであろう。

本稿では潜在クラス分析をもとにした分析を行うことで、データに基づいた分類を目指したい。本稿では若年女性において何種類かの文化的な集団が存在しており、この集団が趣味や関心の領域、そしてネットメディアの利用の選択パターンを規定すると想定する。この選択パターンは社会調査を通じて経験的に観察が可能であり、その選択パターンを潜在クラス分析にかけることで背後に存在する文化的な集団を析出する。ここで先行研究を踏まえた分析上の焦点を4つ挙げておく。

- ① 「高級志向」や「雑食志向」の傾向を持つ集団は、今回のデータ分析で析出されるか。もし析出されるとすれば、社会階層はこれらの文化的分類と関連しているのか。
- ② ファッションや旅行といった領域に趣味・嗜好を持つ集団は、今回のデータ分析で析出されるか。もし析出されるとすれば、この文化的分類はどのような社会属性要因と関連しているのか。
- ③ アニメ、マンガ、ゲームといった領域に趣味・嗜好を持つ集団は、今回のデータ分析で析出されるか。もし析出されるとすれば、この文化的分類はどのような社会属性要因と関連して

いるのか。

- ④ 文化的に活発ではない集団は、データ分析で析出されるのか。もしされるとすればどのような集団はどの程度の割合で存在しているのか。

3. データと手法

3-1. データ

本稿で用いるデータは、2021年1月にA社のオンラインモニターに対して行った「若者の生活に関する調査」である。いわゆる公募型モニターによるウェブ調査であり、20代前半の女性400人、20代後半の女性400人分を回収目標として割り当てたところ、目標数の800を回収することができた。ただし、いわゆる *satisfice* と言われる精度が低いと思われる回答が散見されたため、以下の二段階でデータの選別を行った。第一に、調査項目全体を確認すると、リッカート尺度で特定の数字ばかりを選択しているなどの回答が少なからず見られた。このため、目視による確認により上記のような回答を取り除いた。データクリーニング終了時のサンプルサイズは752であった。第二に、分析に用いる趣味・音楽の好み・ネットメディア利用に関する30項目の合計数の分布を確認したところ、25項目以上回答している者が8名、1項目も回答していない者が31名いることが明らかになった。一般的に考えて上記の項目のうちの大半が当てはまるとは考えがたい一方で、上記のうち1項目も当てはまらないということも考えがたい。前者のようなケースでは数多くの項目にランダムにチェックを入れることで、後者のようなケースではすべての項目に対して「いずれも当てはまらない」という選択肢にチェックを入れることで、*satisfice* を行っている可能性が疑われる。したがって、25項目以上回答している者と1項目も回答していないものをサンプルから取り除い

た。最終的なサンプルサイズは713である。サンプルの選別基準はやや恣意的であるが、精度の低い回答を含めると潜在クラス分析の結果にも影響を与える可能性があるため、ここではこうした回答を取り除いたサンプルを用いる。

分析に用いる項目の内訳としては、「趣味としているもの、または強く興味・関心を持っているもの」14項目、「ふだん聴く音楽ジャンル」9項目、「利用しているネットメディア」7項目である。いずれも当てはまるものをすべて選択する多重回答式であるため、二値変数のデータとして扱われる。今回の調査で趣味・関心の領域だけではなく音楽ジャンルとネットメディアを含めたのは、近年の研究においてもこれらが若年層におけるサブカルチャーの形成と強い関わりを持つものであることが示されており、趣味・関心領域の選択との関連性が期待できるからである（北田・解体研2017; 三浦・天竺2019）。項目の選定にあたっては、北田らの研究を参照しつつ（北田・解体研2017）、2020年代初頭の若年女性にとって一般的であるもの、より具体的には少なくとも5%以上の回答が見込めるものを中心とした⁶⁾。ただし、いわゆる高級文化に属する項目については、若年女性にとってなじみのなさそうなものも加えている（たとえば音楽ジャンルの「クラシック」「ジャズ」など）。

なお、調査の前年の2020年は、新型コロナウイルスの感染が急速に拡大し、室内での活動が社会的に推奨されるいわゆる自粛期間が長く続いた年であった。このため、上記の項目に対する回答はその自粛期間の影響を受けている可能性はあることは付記しておく。

3-2. 潜在クラス分析の推定方法

A. WhiteとT. B. MurphyによるRパッケージBayesLCAの解説論文をもとに、潜在クラス分析のモデルについてまとめておこう（White and Murphy 2014）。まず、M個の項目に関してN人

の回答者がYesとNo、つまり1か0かの二値で回答したとしよう。それぞれの項目に対する回答を確率変数 X と考え、任意の回答者の番号を i 、項目の数を M とする。 i 番目の回答者が m 個目の項目に回答した場合は、 X_{im} と表記することにする。この X_{im} について、 M 個すべての回答を束ねたベクトルを X_i と表記しよう。

ここで、 i 番目の回答者の回答パターン X_i が得られる確率を考えるために、これを $p(X_i)$ と置くことにしよう。この $p(X_i)$ は複数の潜在クラスによって決定され则认为。この潜在クラスの個数をここでは G 個とする。ここで、潜在クラスの商品への回答を決める θ というパラメータを導入しよう。回答者 i の m 番目の回答確率 $p(X_{im})$ が所属するクラスター g のパラメータ θ_{gm} によって決まると考えると、

$$p(X_{im}|\theta_{gm}) = \theta_{gm}^{X_{im}}(1 - \theta_{gm})^{1-X_{im}}$$

である。さらに、回答者 i の M 個の回答すべてに関する $p(X_i)$ を考えると、 $p(X_{im}|\theta_{gm})$ の積を取って

$$p(X_i|\theta_g) = \prod_{m=1}^M p(X_{im}|\theta_{gm})$$

である。

さらに、もう一つパラメータを導入する。それは回答者が各クラスターに所属する事前確率であり、これを τ と置こう（その一つを取り出すときは τ_g と表記される）。 $p(X_i)$ をすべての潜在クラスについて求める場合には、 $p(X_i|\theta_g)$ を計算してそれらを τ_g で個別に重みづけした上で総和を計算する。したがって、以下のように表現できるであろう。

$$p(X_i|\theta, \tau) = \sum_{g=1}^G \tau_g p(X_i|\theta_g)$$

しかし、 θ と τ という2つのパラメータを持つ

上記のモデルでは、解析的に解くことが難しい。これを解決するためにクラスターの所属を示す変数 Z を導入して、この Z を欠測データとみなしてパラメータを推定する。

一般的に用いられるのは最尤推定の一種であるEMアルゴリズムであるが、本稿では事前分布を使用したベイズ推定の一手法であるギブスサンプリングを用いる。ギブスサンプリングはマルコフ連鎖モンテカルロ法（MCMC）のアルゴリズムの一つであり、一定のルールのもとでパラメータの乱数サンプルを大量に生成することで目的の確率分布に近似させることができる手法である。一般的に最尤推定はパラメータの点推定を目的としているが、MCMCを用いたベイズ推定ではパラメータの確率分布を推定できる点に特徴がある。

一回の推定に大量（少なくとも数千回）のサンプリング作業が必要になるため、複雑なモデルを推定するには時間がかかるというデメリットもある。しかしコンピューターやソフトウェアの性能向上により、このようなデメリットもかなりの程度軽減されている。したがって、潜在クラス分析にMCMCを用いたベイズ推定を導入することにも十分な実用性がある。

3-3. ベータ回帰分析の推定方法

次に、推定した潜在クラスと社会属性がどの程度関連しているかという点を推定する方法を考えたい。潜在クラス分析では共変量を投入することで、共変量と潜在クラスの関連を推定することもできる。その一方で、潜在クラス分析で得られたパラメータや予測値を目的変数として、社会属性を投入した回帰モデルを用いることで二段階の推定を行うことも可能である。BayesLCAパッケージでは推定に共変量を利用することができないため、本研究では潜在クラス分析の結果得られる各回答者の潜在クラスへの所属確率を目的変数としたベータ回帰分析を活用する（Ferrari and Cribari-Neto, 2004）。

Rではrstanarmパッケージの関数stan_betaregでは、MCMC、具体的には確率的プログラミング言語Stanによるハミルトニアンモンテカルロ法を用いたベータ回帰分析を行うことができる (Goodrich et al. 2022; Stan Development Team 2023)。ここでAliらの解説を参考にベータ回帰分析の推定方法を確認しよう (Ali et al. 2020)。ベータ回帰分析とは、目的変数の確率分布にベータ分布を想定した回帰モデルの一種である。 $Beta(a, b)$ は0から1の間を取り、2つのパラメータでその形状が決定される。ここで

$$y_i \sim Beta(a, b)$$

とする。ベータ分布の平均 μ と精度（分散の逆数） ϕ については

$$\begin{aligned} a &= \mu \cdot \phi \\ b &= (1 - \mu) \cdot \phi \end{aligned}$$

の関係が成り立つ。ここで、 X を説明変数の行列、リンク関数を g_1 とすると、

$$\mu = g_1^{-1}(X\beta)$$

と表現できる。もし ϕ を定数項とする場合は上記の式のみで推定することになるが、 ϕ を予測する式を立てて組み合わせることも可能である。この場合は、 Z を説明変数の行列、リンク関数を g_2 として、

$$\phi = g_2^{-1}(Z\gamma)$$

となる。上記の式を用いることで、精度が説明変数の大きさによって変化することを想定したモデルを組むことが可能である。なお、目的変数 y_i が0または1を含む場合は、サンプルサイズ N を使った以下の式による変換によって0と1の間の値に

収まるように変換することができる (Smithson and Verkuilen 2006)。

$$y_i' = [y_i(N-1) + 0.5]/N$$

4. 潜在クラス分析の結果

4-1. 事前分布およびMCMCの設定

ギブスサンプリングのようなMCMCでは、事前分布の設定が必要である。潜在クラス分析では潜在クラスごとの各項目の反応率を示す θ にはベータ分布、潜在クラスに対する所属確率 τ にはディリクレ分布を用いる。もしパラメータの分布に関して分析前に使用できる情報があるのであれば、それに基づいて適切な事前分布を与えた方がよい。BayesLCAパッケージのblca関数では、パラメータ θ の事前分布 $Beta(a, \beta)$ と、パラメータ τ の事前分布 $Dirichlet(\delta)$ の事前分布を指定できる。ただし、 τ に関しては集中度パラメータが同じ値になる対称ディリクレ分布のみが指定可能である。

今回の分析ではベータ分布には弱情報事前分布として $\alpha = 1.3$ 、 $\beta = 1.3$ を指定する。この事前分布は、 θ が非常に小さい可能性や非常に大きい可能性は低いと想定していることになる。また、ディリクレ分布のパラメータ τ の集中度パラメータ δ は1.5とする。 δ は1のときに一様分布となり、大きくなればなるほど、それぞれの項目が等確率であるという可能性を高く見積もることになる。今回は1.5という値を設定することで、弱情報事前分布として特定の項目の τ だけが大きいことは可能性として非常に小さいであろうと想定する。

本稿では、これらの事前分布を用いて推定を行うことにしよう。なお、MCMCの設定についてはBayesLCAの初期設定の通り、バーンインを100回、抽出間隔を1、繰り返し回数を5,000回と

表1 潜在クラス分析の結果 (MCMCによる事後分布のMAP推定値)

			Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
		全体	一般趣味	少趣味	アクティブ = オタク	オタク	アクティブ	雑食
			(35.5%)	(25.1%)	(13.0%)	(10.7%)	(10.5%)	(5.3%)
趣味・関心	スポーツ	16.8%	16.3%	10.3%	15.0%	12.2%	40.1%	32.8%
	アニメ	30.0%	8.4%	13.0%	62.7%	86.0%	21.4%	88.6%
	マンガを読む	30.4%	11.2%	10.2%	75.2%	69.6%	20.0%	95.6%
	ゲーム	27.3%	16.4%	8.5%	50.5%	66.2%	20.8%	78.4%
	お笑い	23.0%	16.6%	10.5%	46.5%	16.8%	30.0%	78.0%
	読書	15.6%	4.4%	7.4%	27.6%	28.9%	26.2%	65.8%
	映画・ドラマ	49.1%	48.6%	24.6%	80.0%	40.8%	58.4%	90.7%
	演劇・ミュージカル	13.2%	9.1%	7.8%	10.5%	21.1%	25.0%	48.3%
	カメラ	20.6%	14.0%	6.5%	44.4%	11.6%	38.3%	68.5%
	美術館	14.0%	6.8%	7.7%	11.2%	11.0%	40.4%	67.7%
	旅行	42.9%	44.9%	17.8%	69.8%	23.2%	71.5%	68.8%
	料理	34.4%	29.7%	14.0%	65.0%	31.3%	46.3%	76.2%
	ファッション	45.4%	51.8%	15.3%	74.6%	25.3%	60.6%	86.7%
	ライブ	29.7%	32.8%	9.2%	42.1%	28.4%	43.3%	60.2%
ネットメディア	Instagram	72.5%	88.9%	36.6%	84.2%	56.0%	94.9%	83.3%
	Twitter	64.5%	70.0%	27.8%	78.0%	83.9%	85.0%	82.9%
	pixiv	8.6%	2.5%	4.2%	4.0%	47.0%	5.4%	26.9%
	Facebook	15.0%	14.0%	10.0%	16.2%	7.5%	39.2%	23.1%
	YouTube	77.7%	85.1%	42.0%	96.1%	87.3%	90.2%	96.0%
	LINE	82.5%	96.2%	48.2%	97.0%	82.6%	88.4%	90.5%
	TikTok	16.3%	20.1%	6.0%	21.9%	14.0%	30.5%	15.8%
	ニコニコ動画	10.9%	2.0%	8.5%	2.2%	41.3%	19.5%	41.9%
音楽ジャンル	J-POP	66.3%	74.1%	35.9%	82.8%	70.5%	71.8%	90.6%
	K-POP	26.8%	33.6%	11.4%	37.8%	9.2%	39.2%	47.4%
	洋楽ポップス	23.4%	17.4%	10.4%	31.9%	8.5%	64.4%	63.3%
	ロック	20.3%	10.6%	11.0%	28.2%	30.7%	40.6%	61.3%
	アニソン・声優・ボカロ	24.4%	1.7%	17.0%	21.6%	87.8%	37.3%	75.7%
	ヒップホップ・R&B・レゲエ	14.9%	8.6%	13.4%	9.2%	9.6%	45.6%	42.3%
	ダンス・エレクトロニカ	9.8%	3.8%	7.4%	2.4%	7.0%	39.5%	43.4%
	クラシック	10.7%	3.9%	4.8%	3.4%	18.4%	37.7%	48.5%
	ジャズ	6.9%	1.8%	2.9%	5.0%	7.5%	28.2%	39.4%
	(補足) 平均反応数		9.4	8.5	4.6	13.0	11.4	14.0

注：太字のセルは全体の1.5倍以上、グレーのセルは2/3倍以下の項目選択率を示している

した。また、分析の再現性のためにRの乱数シードを「1」に固定している。なお、トレースプロットやdependence factorを確認した上で、MCMCのサンプリングに問題がありそうな場合には上記の設定を変更してよりよい推定結果を目指した。

4-2. クラス数の検討

潜在クラス分析では潜在クラスの数的分析者が決定する必要がある。潜在クラス分析では、AIC、BIC、DICといった情報量規準を用いたクラス数の決定を行うことがある。ただし、こうした情報量規準を参考にしながらも、理論的な解釈を重視すべきだという考え方もある（Weller et al. 2020）。BayesLCAパッケージではDICが利用できることから、本稿ではDICを参照しながら解釈可能性が最も高いクラスター数を採用したい⁷⁾。

まず、クラスター数の最小を3として、1つずつ増やしながらギブスサンプリングを用いた推定を行った。その際にはMCMCのトレースプロットを参照し、非定常的なサンプリングが行われていないかを確認しながら推定を進めた。その結果として、クラス数が6を超えるとトレースプロットに異常な動きが観察されることが分かった。これはラベルスイッチングと呼ばれるもので、用いたモデルではパラメータが識別不可能であることを示唆している。そこで、潜在クラス数を6～8に設定したモデルに対して、バーンイン回数を1000、抽出間隔を1/4、繰り返し回数を20,000に設定しなおした分析を行ったところ、クラス数が6のモデルについては問題なく収束したが、クラス数が7と8のモデルはやはり収束しなかった。ラベルスイッチングに関しては複雑な要因が絡む問題であるが、7以上のクラスを想定することは今回用いたデータでは適切ではないと判断するのが妥当であろう。したがって、推定された最大のクラス数は6である。

クラス数3から6の推定結果について、情報基準量の一つであるDICを算出した。その結果、ク

ラス数3では-21368.27、クラス数4では-21610.92、クラス数5では-22172.78、クラス数6では-22944.88であり、クラス数6の推定結果のDICが最小であった。また、それぞれの分析結果についての解釈を検討したが、DICが最小のクラス数6でも理論的にも無理のない解釈が可能であった。したがって、今回はクラス数6の結果を採用する。

4-3. 潜在クラス分析の推定結果

表1は、各クラスへの分類確率と各項目の反応率の推定値（各パラメータの事後分布のMAP推定値）を示したものである。また、各クラスの構成比に関しては、MCMCによる事後分布を示す⁸⁾。「全体」という列は、サンプル全体におけるそれぞれの項目への反応率を示しており、その右側に6つのクラスの反応率を示している。分析結果を解釈する際には、各クラスの項目への反応率の大小を確認する必要がある。そこで、今回の結果表ではサンプル全体の項目反応率をもとに、おおよそ目安として1.5倍以上の反応率がある場合は文字を太字に、2/3倍以下の反応率である場合にはセルをグレーとした。あるクラスの反応率が全体の1.5倍を超えていればその項目は全体に比べて多いと解釈し、2/3倍以下であれば少ないと解釈することにしよう。なお、最後の行の「平均反応数」は単純に31項目のパーセンテージを合計して整数にしたものであるが、この値は各クラスの平均反応数を示す（全サンプルの平均は9.4）。

分析結果の解釈は以下の通りである。なお、クラスの番号は予測された分類比率の順である。

1. 一般趣味層（35.5%）。映画・ドラマやファッション、旅行など一般的な趣味には平均的な反応を示すが、アニメやマンガなどのオタク文化に関心が低く、読書や美術館といった教養型の趣味にもあまり関心を示さない層である。ポップス以外の音楽への関心も低い。平均反応数9.4

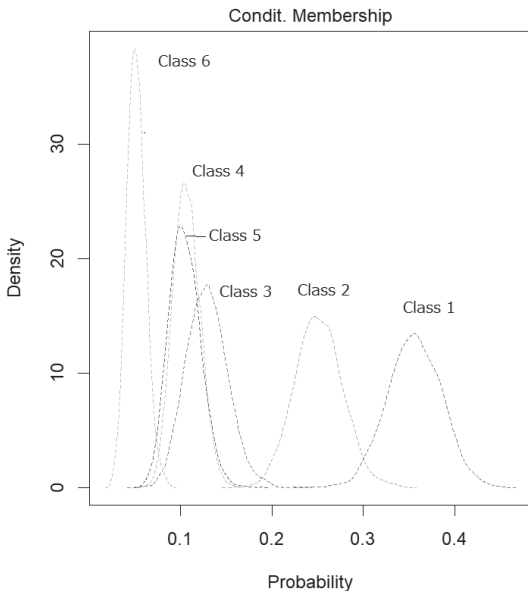


図1 各クラスの構成比 (MCMCによる事後分布)

はサンプル全体よりも少し低い値である。

2. 少趣味層 (25.1%)。ほとんどの項目への反応が低く、平均反応数4.6ということで31項目のうち平均して4～5項目程度しか回答していない層である。全般的に趣味・関心の幅が狭い層であると考えられるが、今回の調査では趣味・関心の領域に対する参与度を問うていないため、この層の中にごく限られた領域に熱中している（そうであるがゆえに他の選択肢を選ばない）人々が混ざっている可能性はある。
3. アクティブ＝オタク層 (13.0%)。映画・ドラマやファッション、旅行などの一般的な趣味にも関心が強いだけでなく、アニメやマンガ、ゲームといったオタク的な趣味にも関心が強い層である。このグループはアクティブな趣味にも関心のあるオタク、ないしはオタク的な趣味を嗜むアクティブ層と解釈することができるであろう。ただし、平均反応数13.0と趣味・関心の幅はやや広い層であるが、音楽ではポップス系以外に対する興味が弱く、pixivやニコ

ニコ動画といったオタク的なメディアへの関心も薄い。

4. オタク層 (10.7%)。アニメ、マンガ、ゲームなどのオタク的な趣味に対する反応率が高い層であり、pixivやニコニコ動画といったオタク系の趣味と親和性の高いメディアの反応率が高く、音楽ではアニソン・声優・ボカロへの反応率が高い。その一方でファッションや海外のポップスへの関心は低いようである。サンプル全体の約1割を占める。
5. アクティブ層 (10.5%)。料理、美術館、演劇・ミュージカル、読書といったインドアの趣味に関心が高いだけでなく、スポーツや旅行といったアウトドアの趣味に関する関心も高い点が特徴的である。また、音楽の各ジャンルに対する関心の高さも特徴的である。また、FacebookやTikTokの利用率が相対的には高い点も特徴的である。平均反応数も14.0とサンプル全体より多い。カテゴリーの名称をつけることは難しいが、スポーツや旅行に対する関心の相対的な高さを重視してアクティブ層と読んでおく。
6. 雑食層 (5.3%)。様々な趣味に関心を示す層である。メディアの利用率も全般的に高く、様々な音楽ジャンルに関心を示す傾向がある。平均反応数は19.8であり、サンプル全体の平均の倍以上である。先行研究で議論されてきた文化的オムニボア（雑食）と呼ばれる傾向を強く示す層である。

以上の分析結果をまとめよう。第一に、全体の6割程度はごく限られた領域にしか関心を示さない層（少趣味層）ないしは一般的な趣味にしか関心を示さない層（一般趣味層）であり、文化的な面で活発と思われるのは残りの4割程度である。第二に、先行研究で議論されてきた「オタク層」「アクティブ層」「雑食層」は潜在クラス分析の結果からもこれに対応するクラスが析出された。第三に、教養のある趣味を志向する「教養層」は本

稿の結果から見出すことはできなかった。

5. ベータ回帰分析の結果

5-1. 事前分布の設定

続いてベータ回帰分析により、社会属性などから前述の潜在クラスへの所属確率を予測するモデルを推定する。ベータ回帰分析においては、回帰係数に対して事前分布を設定する必要がある。今回の分析では回帰係数と切片の事前分布につい

て、`rstanarm`のデフォルト設定である平均0、標準偏差2.5の正規分布である $Normal(0, 2.5)$ を用いる。この事前分布は、分析に用いるすべての説明変数を標準化した場合に、回帰係数と切片が $Normal(0, 2.5)$ に従うことを予想する。具体的には、回帰係数が ± 2.5 の幅を超えることは少なく（確率としてはおよそ32%より小さい）、回帰係数が ± 5 を超えることは稀（確率としてはおよそ5%より小さい）であると考ええる。これは、多数の回帰係数を推定した場合に、0から大きく離れた係数が見られることは稀であり、かつ0から離

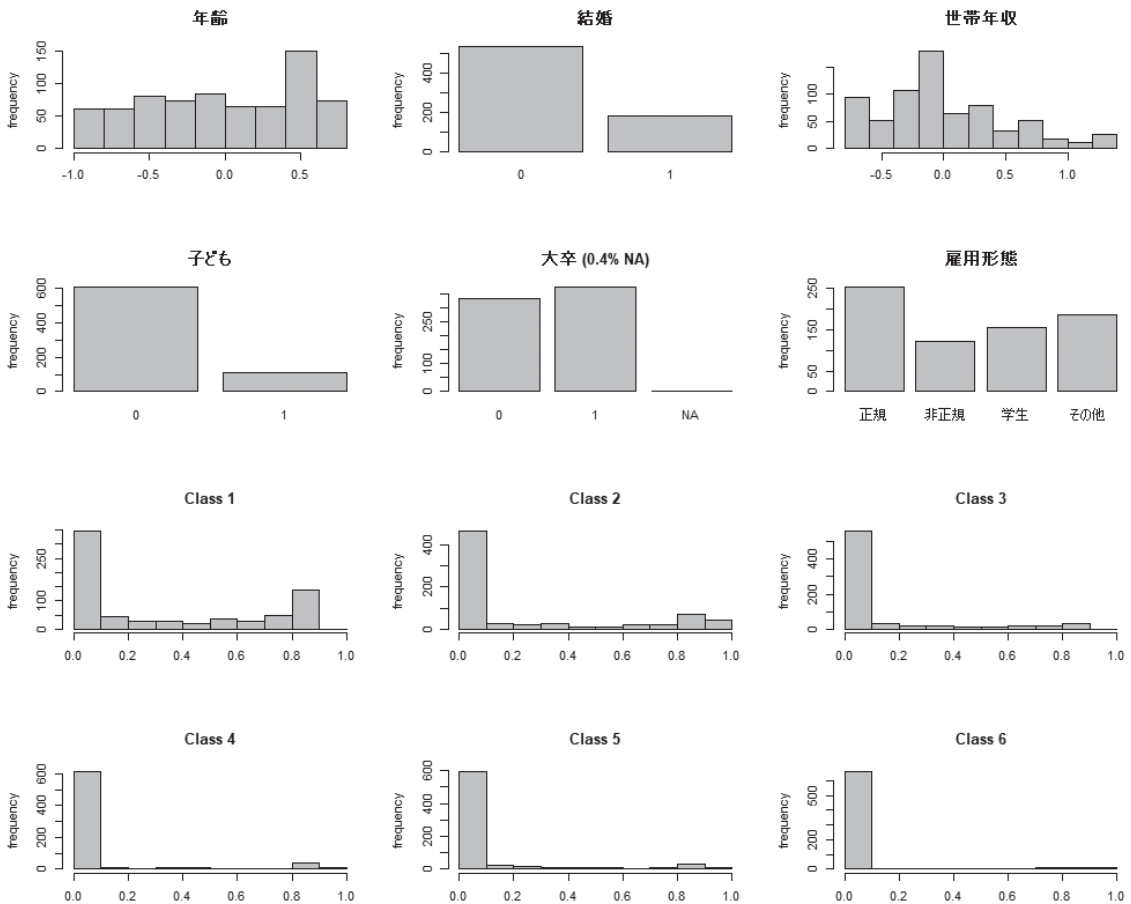


図2 ベータ回帰分析に使用する変数の分布

注：JointAIパッケージ（R）の`plot_all`関数を用いた（Earler et al. 2021）

れた値であればあるほどそのような値が出現する確率は低いであろうという経験的事実を踏まえた弱情報事前分布である。

分析にあたっては、rstanarmのデフォルト設定に従いchainを4回、warmup (burn.inと同意である) を1,000回として、1チェーンにつき1,000回(合計4,000回)の推定を行った。推定の収束についてはRhat値とトレースプロットをもとに判断するものとする。

5-2. 分析に用いる変数

目的変数は6つの潜在クラスに対する所属確率の予測値である。説明変数は年齢(20～29)、結婚(ダミー)、世帯年収(1～15)、子どもあり(ダミー)、大学卒業(ダミー)、雇用形態(正規雇用、非正規雇用、学生、その他の4カテゴリー)である。なお、Gelmanらの推奨に従って、回帰係数の事後分布の範囲を揃えるためにダミー変数以外の変数に対して標準偏差を2倍した値で割って調整した(Gelman et al. 2020)。また、職業ダミーについては正規雇用を参照カテゴリーとするため、正規雇用との対比で解釈することになる。変数の分布は図2に示した通りである。学歴の回答に0.4%の欠測があるが、使用されるのはすべての変数に欠測のない710人のデータである。3-3でも示したように、今回は所属確率の平均 μ と精度 ϕ の双方を推定するモデルを立てるが、分析結果の解釈の中心となるのは平均 μ に関する分析の回帰係数である。

5-2. ベータ回帰分析の推定結果

ウォームアップ回数を1,000回、繰り返し回数を1,000回、チェーンを4回に設定してMCMCに基づく推定を行った。パラメータの推定値のRhat値はすべて1.1を下回っており、トレースプロットにも問題はなかった。したがって、推定は十分に収束していると判断した。

パラメータの事後分布の信用区間(credible

interval、以下CI)は図3に示した通りである。事後分布のCIはパラメータの真値が存在する範囲を示す。今回示したのは、事後分布の中央値に加えて50%(内側)と95%(外側)のCIである。一つの目安としては、95%CIが0をまたいでいなければパラメータが+もしくは-のどちらかに存在する確率はかなり高いと言えるが、ベイズ推定の強みを生かすためには機械的に判断するのではなく、不確実性を連続的に評価することが重要であろう。なお、(phi)と頭についている方が精度 ϕ に関するパラメータであり、(phi)とついていない方が平均 μ に関するパラメータであるが、ここでは μ のパラメータを中心に結果を解釈する。

1. 一般趣味層(Class 1)。年齢(-)および非正規(-)がこのクラスへの所属と関連しているようである。
2. 少趣味層(Class 2)。事後分布の95%CIがプラスまたはマイナスいずれかの領域にのみ存在しているパラメータはなかった。ただし、年齢(+)と結婚(+)はこのクラスへの所属と関連している可能性はあるかもしれない。
3. アクティブ＝オタク層(Class 3)。年齢(-)がこのクラスへの所属と関連しているようである。また、結婚(+), 世帯年収(-), 子どもあり(-), 職業その他(+)がこのクラスへの所属と関連している可能性はあるかもしれない。
4. オタク層(Class 4)。結婚(-), 非正規(+), 職業その他(+)がこのクラスへの所属と関連しているようである。
5. アクティブ層(Class 5)。事後分布の95%CIがプラスまたはマイナスいずれかの領域にのみ存在しているパラメータはなかった。ただし、年齢(+)と結婚(-)はこのクラスへの所属と関連している可能性はあるかもしれない。

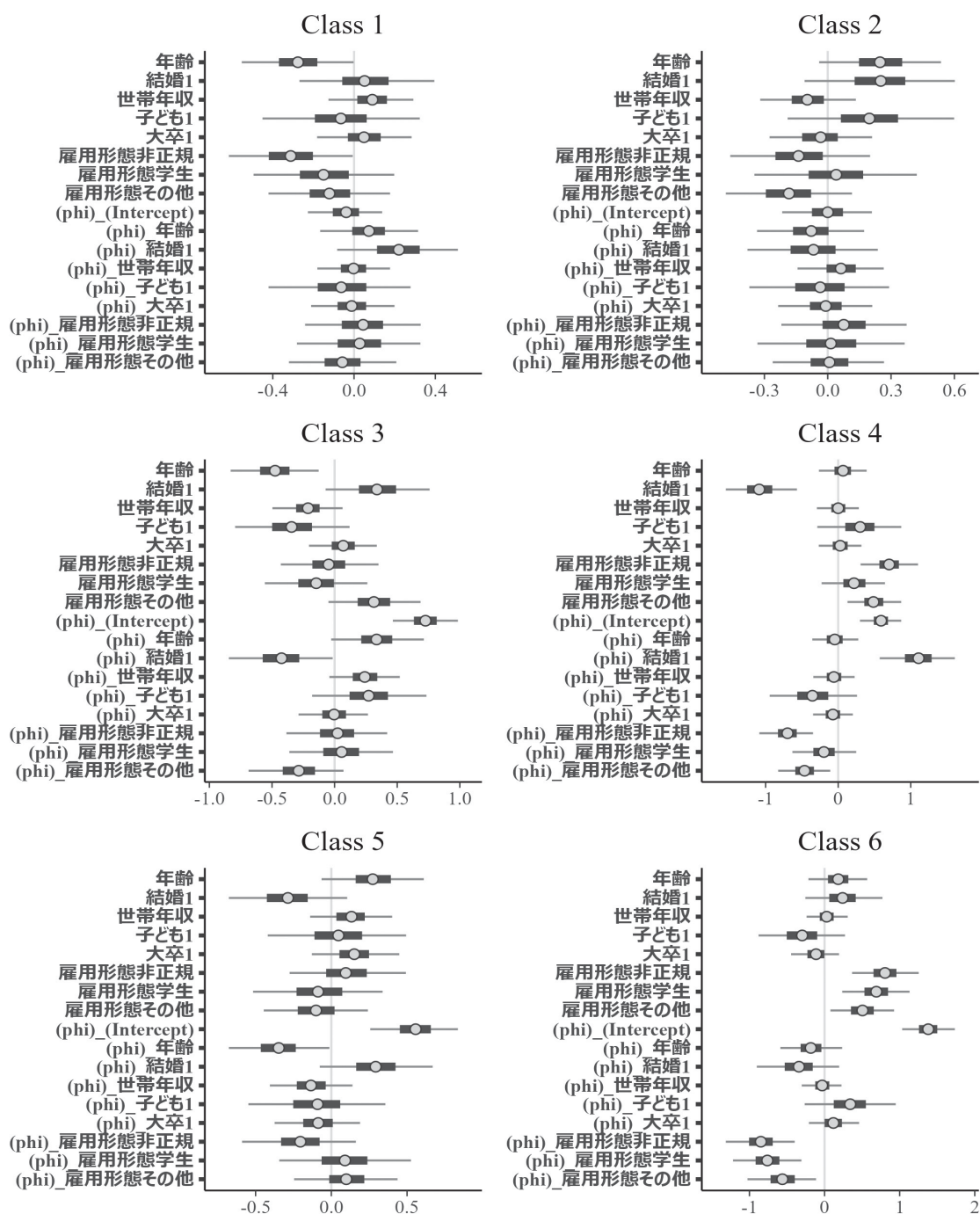


図3 ベータ回帰分析の結果 (MCMCによる事後分布の中央値、50% CI、95% CI)

注：切片の μ に関する点推定値とCIは省略

6. 雑食層 (Class 6)。非正規 (+)、学生 (+)、職業その他 (+) がこのクラスへの所属と関連しているようである。言い方を変えれば、正規雇用はこのクラスに所属する傾向が弱い。

以上の分析結果をまとめると、以下の通りである。

- ・年齢には文化的分類との間に一定の関連があるようである。20代の若年女性の中でもより若い層は一般趣味層やアクティブ＝オタク層に分類されやすい傾向がある。同一とみなされる世代内でも、趣味・関心のあり方に年齢による傾向の違いが表れている可能性がある。
- ・学歴や世帯年収といった社会階層に関連する変数と、今回用いた文化的分類との関連は弱いようであり、先行研究で議論されてきた社会階層と趣味・関心の結びつきは確認できなかった。これは今回の分析対象が若年女性に限定されていることが影響している可能性はあるかもしれない。
- ・結婚・職業と文化的分類との関連をどのように考えるかは難しい問題であるものの、自由に使える余暇時間の多さが介在しているという解釈が妥当であるように思われる。たとえば、オタク層に浸透している「推し活」に必要な時間的資源を考慮すると、既婚者や正規雇用の者が「オタク層」になりづらい（またはもともと「オタク層」であっても結婚や正規雇用職への就職をきっかけに離反しやすい）ことは推測できるかもしれない。

6. おわりに

以上に論じた通り、本稿では20代女性を対象としたウェブ調査のデータから、①趣味・音楽の好み・ネットメディア利用の回答から潜在クラス分析によって若年女性を6つの集団に分類し、②

ベータ回帰分析によって属性的要因と各クラスターへの所属の関連について検討した。最後にいくつかの課題について考察を行いたい。

第一に、本稿が得た6つの潜在クラスについてである。今回得られた6つの潜在クラスは分析者（筆者）の背景知識から見ても無理のない解釈が可能であったため、一般性のある結果であると考えている。しかし、潜在クラス分析を含む混合分布モデルのクラスター（潜在クラス）数の決定は非常に難しい問題であり、議論の分かれるところである。今回用いたデータでは6つというクラスが最適であると判断されたが、サンプルサイズを大きくすればより細かな単位で潜在クラスを析出することもできるかもしれない。したがって、本稿が用いた6つの潜在クラスの一般化可能性については今後検証の余地がある。

第二に、先行研究との整合性についてである。本稿の知見はブルデューの見解や文化的オムニボア仮説とは整合しない部分が少なくなかった。これについては20代の女性のみ調査対象を限定したということも影響しているであろう。ただ、趣味や嗜好に関する先行研究が社会階層の影響に重点を置いてきたため、ライフコースや雇用形態といった余暇活動に大きな影響を与える可能性のある要因にはあまり注目が集まっていなかったようにも思われる。たとえば、正規雇用への就職、結婚、子育てといったライフイベントは、自由に使える余暇時間を抑制すると想定される。一方で、現代社会では趣味や関心を優先させるため、上記のようなライフイベントを回避するといった選択もありうるかもしれない。この点については本稿では十分に理論的な検討を行うことができなかったため、今後の検討課題としたい。

第三に、本研究の知見の一般化可能性について検討したい。公募型モニターによるウェブ調査をもとにした研究であるがゆえに、本研究における知見の一般化には限界がある。また、今回用いた調査では質問項目に多重回答式の選択肢を採用し

たが、江利川・山田（2023）のように多重回答形式は回答者の労力の最小限化（satisfice）を招く可能性があるため個別強制選択形式の方が望ましいという考え方も提示されている。本稿ではデータクリーニングによってsatisficeを行っていると思われる回答者を除外しているため、satisficeの影響は小さくなっていると思われるが、このようなサンプルの選別は恣意的であることも否定できない⁹⁾。回答の傾向から統計的にサンプルを選別する方法の研究も進んでいるため、今後はこのような手法の導入も検討したい。

第四に、本稿では潜在クラス分析とベータ回帰を組み合わせて推定を行った。潜在クラス分析で使ったBayesLCAパッケージでは共変量を設定できないため、二段階に分けた分析を採用したが、二段階の推定ではなく一段階で推定すべきという考え方もあるであろう。この点についても今後の検討課題としたい。

本研究は試論としての性格が強く、データの収集・処理・分析の方法には改善の余地がある。以上に得られた知見をもとに今後さらなる研究の発展を目指したい。

注

- 1) シカゴ学派からバーミングハム学派に至るサブカルチャー概念は、社会規範からの逸脱、社会主流からの周辺化、そして支配的な価値観に対する抵抗といった要素を内包するものであった（Haenfler 2016）。R. Haenflerはこのような観点を受け継ぎ、サブカルチャー研究の視点からゴスやゲーマーといった現代の北米の若者文化を取り上げている（Haenfler 2014）。その一方で、G. A. Fine（1983）やH. Jenkins（2006=2021; 2011）のように、サブカルチャーの研究において逸脱、周辺化、抵抗といった要素をそれほど重視しない論者もあり、本稿ではひとまずこの立場を取ることにしたい。
- 2) 先駆的な研究としては、若年女性のサブカルチャーにおける「ベッドルーム文化」の重要性を指摘したMcRobbie（2000）やアンノン族・コギャルなどを歴史社会学的な視点から分析した難波（2007）などがある。
- 3) 趣味や関心に基づく回答者の分類に潜在クラス分析を適用した例としては、van Reesらの研究の方がChanとGoldthorpeらのプロジェクトに先行していたようである（van Rees et al. 1999）。
- 4) 日本において若年層を趣味・関心に基づいた文化的分

類を行った先駆的な例としては、階層的クラスタリングを用いた宮台ら（2017）がある。ただし、北田ら（2017）は宮台ら（2017）のパーソナリティに関わる要素を含めた類型化を批判的に論じている。

- 5) 個別の趣味領域における事例については、馬場・池田編（2012）や吉光ほか編（2017）を参照されたい。
- 6) 調査項目の選定に際しては、2020年度に筆者のゼミに所属していた学生にも意見を求めており、その結果が一部反映されている。
- 7) DICはDeviance Information Criteriaの略であり、ベイズ推定によるモデル評価にしばしば用いられる。
- 8) 各項目の事後分布をプロットするとグラフが膨大な数になるため、省略する。
- 9) 江利川・山田（2023）の指摘はもっともであるが、個別強制選択形式を採用する場合は回答者の心理的負担感やそれに伴う回答者の脱落なども考慮する必要があるであろう。筆者としては、多重回答形式と個別強制選択形式のどちらを採用するかについては、調査全体の文章量・項目数などを総合的に判断して決定すべき問題であると考えている。

【文献】

- Ali, Imad, Jonah Gabry and Ben Goodrich, 2020, "Modeling Rates/Proportions using Beta Regression with rstanarm," <https://mc-stan.org/rstanarm/articles/betareg.html>
- 浅野智彦, 2011, 『趣味縁からはじまる社会参加（若者の気分）』岩波書店。
- 東園子, 2015, 『宝塚・やおい・愛の読み替え——女性とポピュラーカルチャーの社会学』新曜社。
- 馬場信彦・池田太臣編, 2012, 『「女子」の時代！』青弓社。
- Bourdieu, Pierre, 1979, *La Distinction: Critique Sociale du Judgement*, Paris: Minuit（石井洋二郎訳, 1990, 『ディスタンクシオン——社会的判断力批判 I, II』藤原書店。）
- Chan, Tak Wing and John H. Goldthorpe, 2005, The Social Stratification of Theatre, Dance and Cinema Attendance, *Cultural Trends*, 14(3): 193–212.
- Chan, Tak Wing and John H. Goldthorpe, 2008, Social Stratification and Cultural Consumption: The Visual Arts in England, *Poetics*, 35: 168–190.
- Erler, Nicole S., Dimitris Rizopoulos, and Emmanuel M. E. H. Lesaffre, 2021, "JointAI: Joint Analysis and Imputation of Incomplete Data in R," *Journal of Statistical Software*, 100 (20):1–56.
- 江利川滋・山田一成, 2023, 「複数回答形式と個別強制選択形式の比較」山田一成編『ウェブ調査の基礎——実例で考える設計と管理』誠信書房。
- Ferrari, Silvia and Francisco Cribari-Neto, 2004, "Beta Regression for Modelling Rates and Proportions," *Journal of Applied Statistics*, 31(7): 799–815.
- Fine, Gary Alan, 1983, *Shared Fantasy: Role Playing Games as Social Worlds*, University of Chicago Press.
- 藤原翔・伊藤理史・谷岡謙, 2012, 「潜在クラス分析を用いた計量社会学的アプローチ——地位の非一貫性、格

- 差意識、権威主義的伝統主義を例に」『年報人間科学』33: 43-68.
- Gelman, Andrew, Jennifer Hill, Aki Vehtari, 2020, *Regression and Other Stories*, Cambridge University Press.
- Goodrich B, Gabry J, Ali I, and Brilleman S., 2022, "rstanarm: Bayesian applied regression modeling via Stan." R package version 2.21.3, <https://mc-stan.org/rstanarm/>.
- Haenfler, Ross, 2014, *Subcultures: The Basics*, Routledge.
- Haenfler, Ross, 2016, *Goths, Gamers, and Grrrls*, Oxford University Press.
- Hebdige, Dick, 1979, *Subculture: The Meaning of Style*, Routledge. (山口淑子訳, 1986, 『サブカルチャー——スタイルの意味するもの』未来社)
- 磯直樹・竹ノ下弘久, 2018, 「現代日本の文化資本と階級分化——1995年SSMデータと2015年SSMデータの多重対応分析」石田淳編『2015年SSM調査報告書8 意識I』2015年SSM調査研究会.
- Jenkins, Henry,[1992]2013, *Textual Poachers: Television Fans and Participatory Culture, Updated Twentieth Anniversary Edition*, Routledge.
- Jenkins, Henry, 2006, *Convergence Culture: Where Old and New Media Collide*, New York University Press. (渡部宏樹・北村紗衣・阿部康人訳, 2021, 『コンヴァージェンス・カルチャー——ファンとメディアがつくる参加型文化』晶文社)
- 片岡栄美, 2019, 『趣味の社会学——文化・階層・ジェンダー』青弓社.
- 北田暁大・解体研, 2017, 『社会にとって趣味とは何か——文化社会学の方法規準』河出書房新社.
- 小林盾・大林真也, 2016, 「分析社会学の応用——文化活動はオムニボア(雑食)かユニボア(偏食)か」『理論と方法』31(2): 304-317.
- 近藤博之, 2011, 「社会空間の構造と相同性仮説——日本のデータによるフルデュー理論の検証」『理論と方法』26(1), 161-177.
- 小藪明生・山田真茂留, 2013, 「文化的雑食性の実相——ハイ=ポピュラー間分節の稀薄化」『社会学評論』63(4): 536-551.
- McCutcheon, Allan L., 1987, *Latent Class Analysis*, Sage.
- McRobbie, Angela, 2000, *Feminism and Youth Culture, Second Edition*, Macmillan Press.
- 三浦展・天笠邦一, 2019, 『露出する女子、覗き見る女子——SNSとアプリに現れる新階層』筑摩書房.
- 宮台真司・石原英樹・大塚明子, 2007, 『増補 サブカルチャー神話解体——少女・音楽・マンガ・性の変容と現在』筑摩書房.
- 難波功士, 2007, 『族の系譜学——ユース・サブカルチャーズの戦後史』青弓社.
- Peterson, Richard A., 1992, Understanding Audience Segmentation: From Elite and Mass to Omnivore and Univore, *Poetics*, 21: 243-258.
- Peterson, Richard A. and Roger M. Kern, 1996, Changing Highbrow Taste: From Snob to Omnivore, *American Sociological Review*, 61(5): 900-907.
- Smithson, Michael and Verkuilen, Jay, 2006, A Better Lemon Squeezer? Maximum-likelihood Regression with Beta-distributed Dependent Variables," *Psychological Methods*, 11: 54-71.
- Stan Development Team, 2023, RStan: the R interface to Stan. R package version 2.26.22, <https://mc-stan.org/>.
- 須川亜紀子, 2021, 『2.5次元文化論——舞台・キャラクター・ファンダム』青弓社.
- 田中東子編, 2021, 『ガールズ・メディア・スタディーズ』北樹出版.
- 辻泉, 2015, 「若者たちのパーソナル・ネットワークと「趣味縁」——2007YCRG 杉並調査の結果から」『大妻女子大学人間関係学部紀要 人間関係学研究』17: 145-162.
- van Rees, Kees, Jeroen Vermunt, and Marc Verboord, 1999, "Cultural Classifications under Discussion, Latent Class Analysis of Highbrow and Lowbrow Reading," *Poetics*, 26: 349-365.
- Weller, B. E., Bowen, N. K., & Faubert, S. J., 2020, "Latent Class Analysis: A Guide to Best Practice," *Journal of Black Psychology*, 46(4): 287-311.
- White, Arthur and Thomas Brendan Murphy, 2014, "BayesLCA: An R Package for Bayesian Latent Class Analysis," *Journal of Statistical Software*, 61(13): 1-28.
- 吉光正絵・池田太臣・西原麻里編, 2017, 『ポスト<カワイイ>の文化社会学——女子たちの「新たな楽しみ」を探る』ミネルヴァ書房.

きだ・ゆうすけ/文化情報学部准教授
E-mail : kidayusuke@sugiyama-u.ac.jp