

# 潜在クラス分析の概要とポイント

稲垣佑典

データサイエンス共同利用基盤施設／統計数理研究所

## 1 はじめに

社会科学では異質な特性を有するいくつかの集団の混合によって、母集団が構成されている状態を想定することがままある。そのときには、全体のなかにどういった集団が存在するかを炙り出すことに、研究上の意義を見出すことができるかもしれない。

例えば社会学では、人々の生活における習慣的行動の様式の差異は、彼らが所属する階層によってもたらされているという議論がある (e.g., Bourdieu, 1979=1990)。この議論は、世の人の振る舞い方を場面ごとに切り出せれば、それらを背後の階層構造に応じて幾通りかに分類可能なことを示唆している。そうであれば、行動傾向を尋ねる調査項目への応答パターンをもとに調査対象者を分類して、そこから各階層を類推するという、いわば“逆向き”のアプローチを用いることで、現在の階層構造について効果的な検証ができる可能性が見えてくる。

他にも、市場調査に携わる実務家のなかには、複数の商品の購買データを用いて特定の購買傾向を有する消費者の集団を明らかにしたい、と考えている者も多いだろう。こうした場面において有用な分析手段となるのが、潜在クラス分析である。

潜在クラス分析は Lazarsfeld (1950) が提唱した潜在構造分析のひとつであり、観測変数の反応パターンに基づいて個体をいくつかのグループに分類する方法である。提唱者が

Lazarsfeldであることからわかるように、この分析法は比較的長い歴史的背景をもっている。しかしながら、扱う変数が多数になるとパラメータ推定の計算が複雑になるため、当初は大規模な調査データの分析に用いられることは多くなかった。だが、計算機の演算能力の向上や、手軽に分析を実行できるソフトウェアの登場に伴い、近年では社会科学のみならず様々な分野における適用が拡大している。こうした近況を鑑みて、本稿では潜在クラス分析の概要を知りたい読者に向けた、基本的な解説を行うことにする。

## 2 潜在クラス分析の概要

先述したように潜在クラス分析は、測定可能な変数である「顕在変数 (観測変数)」の関係を、背後の「潜在変数」を考えることで説明する潜在構造分析に分類される技法である。潜在構造分析としてよく知られたものには、因子分析がある。因子分析が顕在変数、潜在変数ともに量的変数を扱うのに対し、今回説明をする潜在クラス分析では質的な顕在変数と潜在変数を扱うという違いがある。

さらに顕在変数について付言すると、潜在クラス分析では名義的な応答としての顕在変数の他に、各応答の間に順序関係が想定される顕在変数も使用できる。ただし、そこから導出される潜在変数の部分集団 (クラス) は独立したのものとして定義されるので、順序性はもたない。よって潜在クラス分析は、質的な変数を多用す



る社会調査と親和性が高く、結果の解釈に関してもクラス同士の関連性を意識する必要がないため、結果についてシンプルな議論をしやすい分析法であるといえる。

部分集団を想定しない因子分析では、データを構成する全ての個人が同一の測定モデルに従うと仮定され、因子によって個人を分類することよりも、類似した応答パターンをもつ変数間の関係を明らかにすることに重きが置かれる。一方、潜在クラス分析では、個人の応答パターンの類似性からクラスが分類される。そして、所与のクラスである顕在変数に対して特定の応答がなされる確率、つまり条件付き応答確率からクラスの特徴を読み取る作業を通じて、大きな集団を構成する部分集団の異質性を解き明かす。

このように、因子分析と潜在クラス分析は潜在構造分析という括りで扱われているものの、重視している点には多少なりとも差異がある。また、そうした性質の違いのことを「変数志向的アプローチ (Variable-Oriented/Cantered Approach)」、 「個人志向的アプローチ (Person-Oriented/Cantered Approach)」と呼称する場合もある (Collins and Lanza, 2010; 藤原・伊藤・谷岡, 2012)。

### 3 潜在クラス分析の数理的要点

つづいて潜在クラス分析の数理的な要点を簡潔に述べる。ここでは稲垣・朴・前田・中村 (2017) での潜在クラス分析の記述をベースに、McCutcheon (1987) および渡辺 (2001) の説明を適宜援用ながら解説する。さらに技法の詳細を知りたい場合は、Collins and Lanza (2010), Hagenaars and McCutcheon (2002) などを別途ご覧いただきたい。また対数線形モデルに馴染みのある方であれば、潜在クラス分析を対数線形モデルの拡張として説明している三輪 (2009), 藤原・伊藤・谷岡 (2012) を

参照されると、一層理解が深まるだろう。

まず  $A, B, C$  という3つの質的な顕在変数があるとしよう。この顕在変数は、社会調査における質的な質問項目であると考えてもらえばよい。また、それぞれの顕在変数には  $A_i (i = 1, \dots, I), B_j (j = 1, \dots, J), C_k (k = 1, \dots, K)$  という、 $I, J, K$  種類の離散的な状態の応答があるとするとする。これらは、「あてはまる／あてはまらない」のような回答選択肢のカテゴリに相当する。

さらに、その背後には  $A, B, C$  への反応を規定する  $X$  という観測不能な一つの潜在変数があり、それは  $X_t (t = 1, \dots, T)$  となる  $T$  個の状態をとるとする。この  $T$  個の状態をとる潜在変数が「潜在クラス」であり、これがいくつに分けられるのかによってクラス数が決まり、それぞれのクラスで各変数の応答確率がどうなるかを推定するのが、潜在クラス分析の骨子である。

なお、各変数の応答確率については、次のように定義できる。

$\pi_t^X$ : 潜在変数  $X$  に対して個人が潜在クラス ( $t$ ) に帰属する確率 (潜在確率)

$\pi_{ijk}^{ABC}$ : 顕在変数  $A, B, C$  に対して ( $ijk$ ) となる同時確率 (顕在確率)

$\pi_{ijkt}^{ABCX}$ : 潜在変数  $X$  を仮定したときに  $A, B, C, X$  に対して ( $ijkt$ ) となる同時確率

$\bar{\pi}_{ijkt}^{ABCX}$ : 潜在変数  $X$  の潜在状態が所与のもとで顕在変数  $A, B, C$  に対して ( $ijk$ ) となる条件付き確率 (変数の上にバーを付けた  $\bar{\pi}$  で条件付き確率を表記)

その際、潜在変数  $X$  における  $T$  個の状態をとる潜在クラスへの帰属確率の総和を1とする制約が置かれる。これにより、潜在変数  $X$  には互いに背反な潜在クラスが存在し、各個体はその

中の一つのクラスに属することになる。

$$\sum_{t=1}^T \pi_t^X = 1$$

潜在クラス分析では、各個人の潜在変数  $X$  における帰属状態を含む同時確率と、観察された  $\pi_{ijk}^{ABC}$  との間に、次のような混合分布モデルを仮定する：

$$\pi_{ijk}^{ABC} = \sum_{t=1}^T \pi_{ijkt}^{ABCX}$$

これはすなわち、観測可能な確率  $\pi_{ijk}^{ABC}$  が、ある個人が  $(ijkt)$  という特定の反応をする確率  $\pi_{ijkt}^{ABCX}$  の、潜在変数  $X$  中  $T$  個の潜在クラスにおける応答確率の和となることを表している。

このときに  $\pi_{ijkt}^{ABCX}$  は、条件付き確率を使えば

$$\pi_{ijkt}^{ABCX} = \pi_{ijkt}^{\overline{ABC}CX} \pi_t^X \quad (1)$$

となり、 $\pi_{ijkt}^{\overline{ABC}CX} \pi_t^X$  については次のようにモデル化する：

$$\pi_{ijkt}^{\overline{ABC}CX} \pi_t^X = \pi_{it}^{\overline{A}X} \pi_{jt}^{\overline{B}X} \pi_{kt}^{\overline{C}X} \pi_t^X \quad (2)$$

モデル式 (1) では  $\pi_{ijkt}^{ABCX}$  が、潜在変数  $X$  を所与としたときに個人が  $(ijk)$  という反応をする条件付き確率  $\pi_{ijkt}^{\overline{ABC}CX}$  と、当該個人が潜在変数  $X$  に対して  $(t)$  へ帰属することを示した確率  $\pi_t^X$  の積で表されている。

(2) の右辺は、 $X$  が与えられたもとで個人が各顕在変数  $A, B, C$  に対応した  $(i), (j), (k)$  に応答することを示す条件付き確率  $\pi_{it}^{\overline{A}X}, \pi_{jt}^{\overline{B}X}, \pi_{kt}^{\overline{C}X}$  の積として、 $\pi_{ijkt}^{\overline{ABC}CX}$  が表現されることを意味している。これは  $X$  所与のもとでの変数  $A, B, C$  に対する応答の独立性を表しており、「局所独立の仮定」と呼ばれる。

この局所独立の仮定は、ある顕在変数（仮に  $y_1$  と  $y_2$  とする）間の連関が、潜在変数  $X$  の統制によって消失することを指す（直感的なイメー

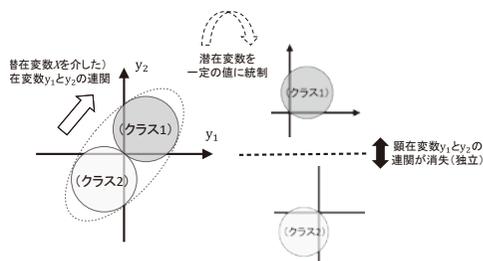


図1 「局所独立の仮定」イメージ図

ジとしては図1のようになる。ただし、当該図は量的変数の散布図に相当し、質的変数の連関をそのまま表すものでないことに注意されたい。先に触れた潜在クラス分析でのクラス数の決定とは、いくつかのクラスを考えればこうした局所独立の状態に至るのかを、理論的仮説あるいは統計的基準を用いて探り出すことに該当する。

さらに既述した  $A, B, C$  の応答に対しては、潜在変数  $X$  の  $(t)$  が与えられたもとでの条件付き確率を与えることに付随して、以下のような確率の総和が1となる制約を伴う。

$$\sum_{i=1}^I \pi_{it}^{\overline{A}X} = \sum_{j=1}^J \pi_{jt}^{\overline{B}X} = \sum_{k=1}^K \pi_{kt}^{\overline{C}X} = 1$$

これら条件付き確率  $\pi_{it}^{\overline{A}X}, \pi_{jt}^{\overline{B}X}, \pi_{kt}^{\overline{C}X}$  をクラス  $1, 2, \dots, T$  ごとにまとめたものを、潜在クラスのプロファイルと呼ぶ。潜在クラス分析では、このプロファイルを用いて抽出されたクラスが何を示しているか解釈を加える。ちなみに、後に出てくる表3（潜在クラスプロファイルの一覧）で、ここまでの説明における記号  $X$  にあたるのが、クラス1, クラス2という潜在クラスである。また、記号  $A, B, C$  は、同じく後述する  $hi\_mot, hi\_read, \dots$  といった変数に対応している。同様に  $(i), (j), (k)$  は、各変数の応答カテゴリである、「非該当」／「該当」にあたる。さらに、上式における  $\pi_{it}^{\overline{A}X}$  の部分によって、所与のクラスにおいて、ある変数に対して「該当」する（または「非該当」である）との応答がなされる確率が表さ



れている。

そして、実際に生じた個々の応答パターンの度数 ( $n_{ijk}^{ABC}$ ) (顕在確率としては  $\pi_{ijk}^{ABC} = \frac{n_{ijk}^{ABC}}{N}$ ,  $N$  はサンプルサイズ) の情報を利用して、潜在変数から想定される母集団の構造が推測される。例えば最尤推定において、セル( $ijk$ )の観測度数  $n_{ijk}^{ABC}$  について多項分布を仮定して導かれる対数尤度関数は (定数項を無視すれば), 本質的に次の形をもつ:

$$L = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K n_{ijk}^{ABC} \log \pi_{ijk}^{ABC} \\ = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K n_{ijk}^{ABC} \log \left[ \sum_{t=1}^T \pi_{it}^{AX} \pi_{jt}^{BX} \pi_{kt}^{CX} \pi_t^X \right] \quad (3)$$

なお、ここでの対数尤度  $L$  を最大化するパラメータの推定には、潜在クラスへの帰属を欠測データと捉えたEMアルゴリズム (詳細は Dempster, Laird and Rubin (1977) などを参照のこと) を用いるのが一般的である。

以上が基本的な潜在クラス分析の数理的要点であるが、潜在クラスを得るために用いた顕在変数のみならず、その潜在クラスの説明要因となる顕在変数を共変量として投入して、潜在クラスへの効果を推定することも可能である (潜在クラス多項ロジスティック回帰分析などと呼称される)。他にも地域や時間で入れ子状になったデータを用いて、マルチレベル分析や多

母集団同時分析, 経時的な分析 (Latent Class Growth Analysis (LCGA), Latent Transition Analysis (LTA) など) を行うこともできる。

## 4 分析にあたってのポイント

ここからは、潜在クラス分析を実施するうえで押さえておくべき事柄を、具体例を示しながら述べる。例示に用いたデータは、UCLAのサイト (<https://stats.idre.ucla.edu/mplus/code/a-latent-class-example/>) からダウンロードできるので、練習の際に活用するとよいだろう。

### 1 分析データと変数

表1は、今回使用する分析データの一部を示したものである。これは (架空の) 生徒の学業成績に関連した質的変数の一覧表であり、1つ目の hi\_mot は「意欲の得点が中央値以上か (意欲\_高)」、2つ目の hi\_read は「読解の得点が中央値以上か (読解\_高)」、3つ目の hi\_math は「数学の得点が中央値以上か (数学\_高)」、4つ目の hi\_sci は「理科の得点が中央値以上か (理科\_高)」、5つ目の hi\_ss が「社会の得点が中央値以上か (社会\_高)」を表している。

例示に用いたデータのサンプルサイズは600であるが、Finch and Bronk (2011) では、潜在クラス分析を行う際の最低限のサンプルサイズ

表1 分析データにおける変数と変数の値

ID	hi_mot (意欲_高)	hi_read (読解_高)	hi_math (数学_高)	hi_sci (理科_高)	hi_ss (社会_高)
1	2	1	1	1	2
2	2	1	1	1	1
3	1	1	1	1	1
4	2	1	1	1	1
5	1	1	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮
596	1	2	2	2	2
597	1	2	2	2	2
598	2	2	2	2	2
599	2	2	2	2	2
600	1	2	2	2	2

$N=600$

数値コード: 1 = 非該当, 2 = 該当

を500以上としている。サンプルサイズが小さいと推定結果が安定せず、クラスの分類がうまくいかないことがあるため、できれば大きなサンプルサイズのデータを用いたいところである。

今回は全てが非該当／該当という2値をとる質的変数を用いたが、3値以上をとる変数や、応答の間に順序性のある変数であっても分析できる。とはいえ、多くの値をとる複数の変数を投入するとクラスの解釈が難しくなるだけでなく、推定にも時間がかかるので注意して欲しい。

また、潜在変数との連関が薄い変数を用いると、モデル適合度が低くなったり、うまくクラスが推定できなかつたりという問題が生じやすくなる。そのような変数は確認の後に、分析から除外するとよい(渡辺, 2001)。逆にいえば、Wurpts and Geiser (2014) で示されたように、潜在変数と高い連関のある“質の良い”変数を見つけれれば、多少サンプルサイズが小さくとも適合的な結果が得られやすくなるということである。

なお、オリジナルのデータは、いずれも非該当／該当が0/1で表されたダミー変数であった。だが、一部ソフトウェアでは、値が「1」以上でないと受け付けてくれない(例えばWinLTAでは「0」は欠損値扱い)。そのため、ここでは非該当=1、該当=2となるように変換した。このように、ソフトウェアの運用を視野に入れた変数操作が必要となる場合もある<sup>1)</sup>。

## 2 モデル選択によるクラス数の決定

潜在クラス分析では、はじめに適合度指標や尤度比検定の結果に基づき、いくつかのクラスのモデルを採用するか選択しなければならない。

モデル選択時の適合度指標には、AICとBICを用いるのが一般的である。

尤度比検定では通常の尤度比 $\chi^2$ 検定に加えて、ブートストラップ法による尤度比検定(BLRT)が使用されることもある。とはいえ、全てのソフトウェアにBLRTが実装されている訳ではなく、依然としてAIC, BICによるモデル選択が行なわれることも多い。(Nylund, Asparouhov and Muthén (2007) はBLRTを推奨しているが、次善の方法としてBIC, ABICを挙げている。)

表2には、分析したモデルごとの適合度指標が示されている。今回は一般的な指標であるAIC, BICを用いて3クラスまでのモデルを推定したが、2クラスモデルがAIC, BICともに適合的な値となっているのがわかる。また、導出されたクラスに対して、サンプルが適切に分類されたことを表す指標であるEntropy ( $R^2$ のように0~1の値をとり、0.80以上が正確とされる(Clark and Muthén, 2010))も、モデル2では0.89という十分に許容できる値であった。よって、2クラスモデルが適当であると判断できる。

モデル選択時には適合度、分類の正確性指標の他にも、次節に示す潜在クラスのプロファイルも参照しておくべきである。もしそこで変数の条件付き応答確率に、「0」や「1」という境界値が多数見られたら、複雑すぎるモデルへの過剰適合が生じており、実態が正しく結果に反映されていないおそれがあることに注意してほしい。また、そういったモデルでは、推定時の計算に長い時間がかかることがある。クラス数が増えれば推定の時間は伸びるのが常であるが、それでも計算に非常に長い時間を要し、クラスプロファイル中に境界値が頻出しているならば過剰適合を疑い、そのモデルは採用しな

表2 モデル選択のための指標一覧

モデル	対数尤度	AIC	BIC	Entropy
1クラスモデル	-2040.12	4090.25	4112.23	1.00
2クラスモデル	-1789.70	3601.40	3649.76	0.89
3クラスモデル	-1785.57	3605.14	3679.89	0.91



い方が無難である。このように潜在クラス分析では、適合度指標、分類の正確性、クラスプロフィールの状態（および解釈可能性）、事前の仮説との対応関係などを総合的に考慮して、モデルを選択しなければならない。

さらに、推定にEMアルゴリズムを用いるソフトウェアでは、初期値に依存して推定値が本来の最適解ではない局所解に陥ってしまうことがある。この問題への対策として、一部のソフトでは異なる初期値を生成して反復計算をしてくれるが、そうした機能を有しないものもある。その場合は、手動か外部のプログラムを用いて同じモデルを繰り返し計算し、各指標や潜在クラスプロフィールに齟齬が生じていないか確認をする必要がある。

### 3 クラスの解釈

いくつかのクラスをもつモデルを採用するか決定したら、続いて潜在クラスのプロフィールを参照しながら、それぞれのクラスが何を表しているか解釈する。その作業は、因子分析で因子負荷量をもとにして因子に対する解釈を与えるプロセスと似ている。

表3には、潜在クラスプロフィールと変数の情報がまとめられている。この表を見ると、クラス1ではhi\_read, hi\_math, hi\_sci, hi\_ssに対して、「該当」と応答される確率が一様に高いこ

とがわかる。これらの変数は、主要な教科の得点が中央値以上であることを表しているため、このクラスに属する生徒の学校での成績は概ね良好であることがうかがえる。けれども、同クラスにおけるhi\_motへの条件付き応答確率は、該当と非該当で同程度であったことから、必ずしも意欲が高い生徒ばかりであるとはいえない。他方、クラス2ではクラス1とは対照的に、hi\_read, hi\_math, hi\_sci, hi\_ssにおいて「非該当」と応答される確率が高く出ており、学業成績はさほど良好ではないようだ。そのうえ、クラス2ではhi\_motに対して、「非該当」という応答がなされる確率が高いことから、意欲の低さも垣間見える。

以上をまとめると、クラス1については、意欲はともかく主要教科の成績は良い、「成績良好」のクラスであると解釈できる。クラス2は、学びへの意欲が欠如気味で成績もあまり芳しくない、「学業低迷」クラスであると解釈できるだろう。これらの結果を俯瞰すると、成績が良い生徒の意欲が高いとは限らないが、成績が低迷している生徒は同時に意欲も低下した状態にあるという興味深い傾向も見えてくる。その他にクラス構成割合に目を向けると、「成績良好」のクラス1は全体の約57%を占めており、クラス2の「学業低迷」は44%ほど存在していることがわかる。

表3 生徒の学業成績に関する潜在クラスプロフィール

顕変数	応答カテゴリ	潜在クラス	クラス1	クラス2
		クラス構成割合 (N=600)	(成績良好)	(学業低迷)
hi_mot(意欲_高)	非該当(n=357)		0.511	0.705
	該当(n=243)		0.490	0.295
hi_read(読解_高)	非該当(n=285)		0.071	0.999
	該当(n=315)		0.929	0.001
hi_math(数学_高)	非該当(n=302)		0.245	0.838
	該当(n=298)		0.755	0.162
hi_sci(理科_高)	非該当(n=284)		0.195	0.835
	該当(n=316)		0.805	0.165
hi_ss(社会_高)	非該当(n=211)		0.142	0.624
	該当(n=389)		0.858	0.376

なお、多数の潜在クラスがあり、変数の数や応答カテゴリの数も多い状況では、潜在クラスプロファイルの情報が複雑になりがちである。その場合、表だけを用いてクラスを解釈しようとすると、かなりの苦勞を要する。そうしたときには、図2のように条件付き応答確率を図としてプロットすると、クラスの特徴を理解するうえでの助けとなる。

#### 4 分析ソフト

潜在クラス分析を実行するためのソフトウェアにも簡単に触れておく。

商用のものとしては、Latent Gold, M-plus, SAS (および JMP), SPSS Amos, Stata<sup>2)</sup>, WIN-MIRA 2001などがある。やや特殊なものでは、Latent GoldをもとにExcelへの機能追加プログラムとして開発されたExcelアドイン潜在クラス分析Ver.1.0やXLSTATがある。フリーソフトには、R (poLCAが代表的だが、潜在クラス分析が実行可能なパッケージは複数ある), LEM, WinLTA, HADなどがある<sup>3)</sup>。商用、フリーソフトいずれも十分な機能を備えているが、使用さ

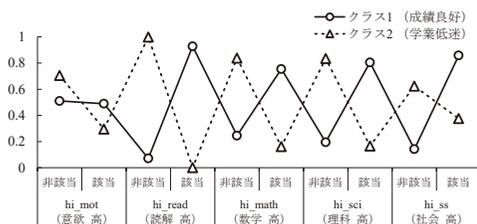


図2 条件付き応答確率のプロット

れている推定法や算出される統計量の種類には多少の違いがある。他にもGUIの有無、読み込み可能なデータファイルの形式など、各ソフトウェアを特徴づける様々な機能があるので、それらを調べて各自の目的・状況に適したものを使用するとよい。

## 5 おわりに

本稿では、社会調査データを分析するうえで有用性が認識され、広く用いられるようになった、潜在クラス分析の基本的な考え方と分析のポイントを紹介した。ここで述べた事柄は初歩的なものであり、現在はより洗練されたモデルも数多く開発されている。それらについて知りたい場合は、先に挙げたCollins and Lanza (2010)などを参照いただきたい。

潜在クラス分析は、単に個人を集団へと分類するためだけでなく、多義性を有した概念の実質的な意味を探るためであったり、異なる特性をもった複数のサブグループの構成状態から社会全体のありさまを浮き彫りにしたりと、我々に様々な分析可能性をもたらしてくれる有用な分析技法である。この技法によって、今後より多くの研究が発展することを期待したい。

#### 謝辞

本稿の執筆に際し、データサイエンス共同利用基盤施設/統計数理研究所 前田忠彦准教授より有益なコメントをいただいた。ここに記して感謝の意を表す。



## 注

- 1) 本稿ではデータ分析にLatent Goldを使用しているため、本来は数値変換不要である。
- 2) Stataはバージョン15から公式に潜在クラス分析が実装された。なお、バージョン11以降であれば、外部プラグイン (<https://www.methodology.psu.edu/downloads/lcastata/>) を利用して潜在クラス分析を実行できる。
- 3) 各種フリーソフトへのリンクは以下を参照のこと。
  - ・poLCA (Rパッケージ) : <https://cran.r-project.org/web/packages/poLCA/index.html>
  - ・LEM (Jeroen K. Vermunt氏 (Tilburg University) が開発) : <https://jeroenvermunt.nl/lemfiles/lemfiles.zip>
  - ・WinLTA (Penn State's College of Health and Human Development Methodology Center) : <https://www.methodology.psu.edu/downloads/winlta/>
  - ・HAD (清水裕士氏 (関西学院大学) が開発) : <http://goo.gl/3FlmBW>

## 文献

- Bourdieu, P., 1979, *La distinction: Critique sociale du jugement*, Paris: Éditions de Minuit. (石井洋二郎訳, 1990, 『ディスタクシオン I・II——社会的判断力批判』藤原書店.)
- Collins, L. M. and S. T. Lanza, 2010, *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Hoboken, NJ: John Wiley & Sons, Inc.
- Clark S. L. and B. Muthén, 2009 “Relating Latent Class Analysis Results to Variables Not Included in the Analysis.” (<http://www.statmodel.com/download/relatinglca.pdf>) (2019年12月11日アクセス).
- Dempster, A. P., N. M. Laird and D. B. Rubin, 1977, “Maximum Likelihood from Incomplete Data via the EM Algorithm.” *Journal of the Royal Statistical Society* 39(1): 1-38.
- Finch, W. H., and K. C. Bronk, 2011. “Conducting confirmatory latent class analysis using Mplus.” *Structural Equation Modeling* 18(1): 132-151.
- 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.
- Hagenaars, J. A. and A. L. McCutcheon (Eds.), 2002, *Applied Latent Class Analysis*. New York, NY: Cambridge University Press.
- 稲垣佑典・朴堯星・前田忠彦・中村隆, 2017, 「地域活動不参加者のプロファイリングは可能か?——地域活動不参加要因の類型化に基づく背景分析」『計画行政』40(3): 33-43.
- Lazarsfeld, P. F., 1950, “The Logical and Mathematical Foundation of Latent Structure Analysis.” In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, and J. A. Clausen, *Measurement and Prediction: Studies in Social Psychology in World War II*. 4: 312-361. Princeton, NJ: Princeton: Princeton University Press.
- McCutcheon, A. L., 1987. *Latent Class Analysis*. Newbury Park, CA: Sage.
- Nylund, K. L., T. Asparouhov and B. O. Muthén, 2007, “Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study.” *Structural Equation Modeling* 14(4): 535-569.
- 三輪哲, 2009, 「潜在クラスモデル入門」『理論と方法』24(2): 345-356.
- 渡辺美智子, 2001, 「因果関係と構造を把握するための統計手法——潜在クラス分析法」岡太彬訓・木島正明・守口剛編『マーケティングの数理モデル』朝倉書店.
- Wurpts, I. C., and C. Geiser, 2014, “Is adding more indicators to a latent class analysis beneficial or detrimental?: Results of a Monte-Carlo study.” *Frontiers in Psychology* 5: article 920. (<https://doi.org/10.3389/fpsyg.2014.00920>) (2019年12月11日アクセス).