

# エビデンスに基づく教育のための縦断データの解析方法 Longitudinal Data Analysis Methods for Evidence-Based Education

玉井 航太 TAMAI, Kota

● 北海商科大学  
Hokkai School of Commerce

藤田 英典 FUJITA, Hidenori

● 共栄大学  
Kyohei University

 **Keywords** エビデンスに基づく教育, RCT, 縦断データ  
evidence-based education, RCT, longitudinal data

## ABSTRACT

近年、「エビデンスに基づく教育」が注目を集めている。「エビデンス」という言葉は実証された科学的根拠の意味で用いられ、ランダム化比較試験 (RCT) による実験デザインを用いた研究が焦点化される。しかし、教育という社会的現実が流動的であり、多くの変数で構成されるため、RCTによる実験デザインは因果関係の検証に有効だが、社会調査の文脈において現実的ではない。そこで、本稿はRCTの方法論的側面における既存の問題を議論し、縦断データの解析方法を紹介した。まず、教育の文脈におけるRCTに関する方法論的問題を概観した。次に、2時点データから因果関係を分析するモデルとして交差遅延効果モデルと同時効果モデルについて述べ、3時点以上のデータを分析するモデルとして、潜在曲線分析、潜在曲線混合モデル、混合軌跡モデルの基本的概念を説明した。最後に、本稿は、教育学的研究が実証科学のための適切な方法論を用いることを提唱し、エビデンスに基づく教育のための方向性を示唆した。

Recently the term “evidence-based education” has been attracting a great deal of attention. The word ‘evidence’ is used to mean substantiated scientific grounds, based on studies using experimental designs that are known to gather data through randomized control trials (RCT). However, while an experimental design

with RCT is effective for examining causal relationship, it is not realistic in the context of social surveys, because the real social world of an educational setting is highly fluid, confining many variables. Thus, in this paper the existing problems with RCT in terms of methodological aspects are discussed, and analytical methods for longitudinal data are introduced. First, this paper examines methodological issues concerning RCT in the context of education. Second, a cross-lagged effects model and a synchronous effects model are described as models by which to analyze causal relationship in data collected over two points in time. Third, basic concepts of latent curve analysis, a latent curve mixture model, and group-based trajectory analysis are explained as a model by which to analyze data collected over three or more points in time. Finally, this paper proposes that educational research use an appropriate methodology for empirical science, and suggests directions for evidence-based education research.

## はじめに

教育分野において「エビデンスに基づく教育」という用語が広まっている。この「エビデンスに基づく」という考え方は、「エビデンスに基づく医療（Evidence Based Medicine：以下EBM）」という考え方から派生したものである。そして、1990年代後半以降、教育を含めた様々な分野でエビデンスの概念が広まるようになった（津谷、2000）。エビデンスに基づく教育という考え方は重要ではあるが、教育分野において、EBMが重視するランダム化比較試験（Randomized Controlled Trial：以下RCT）を用いた研究だけをエビデンスとして扱うことによって限界も生じる。そこで、本稿は、教育分野により適合したエビデンスを生み出す可能性を持つ縦断データについての代表的な分析手法を概観しつつ、「エビデンスに基づく教育」という文脈における縦断的データの有効性について検討する。

## 1. 教育分野におけるRCTの限界

EBMではエビデンスの質に差が存在しており、エビデンスとしての検証方法によって、その信頼性の程度が異なる（津富、2003）。最も良いエビデンスとみなされるのは、RCTに基づいた研究の系統的レビューであり、最もエビデンスの質が低いものとして専門家の意見が位置付けられている。RCTとは、介入群と対照群に対して被験者をランダムに割り当て、その2つの群の違いを統計

的に比較するというのが基本的な在り方である。RCTに基づく知見が強いエビデンスとして成立するのは、因果関係を強く検証することができるという点にある。ランダム割り当てがなされた上での実験デザインであるからこそ、恒常化、均衡化、相殺などの統制法によって原因から結果に至るまでの剰余変数の影響を除去することができる。その結果、独立変数と従属変数の因果関係以外の説明可能性が排除されることになり、介入とその結果としての効果という因果関係を特定することができる。そして、そのようなRCTによる研究結果の系統的レビュー、つまり、メタアナリシスによって、文脈や標本の違いを超えて独立変数と従属変数の因果関係を明確化することができ、効果を評価することができるのである。

しかしながら、教育分野ではRCTに基づく質の高いエビデンスが産出されづらいというのが現実である（Fitz-Gibbon, 2000）。EBMの世界的な波及に基づき、社会科学分野でもキャンベル共同計画が2000年に発足し、エビデンスに基づくことの普及が行われた。しかし、教育分野において、RCTをベースにしたエビデンスは少ない（岩崎、2010）。この背景には現実として多くの制約がRCTに課せられるということがあるだろう。岩崎（2010）は自らが実施した小学校での実験研究をRCTの観点に照らし合わせて検討し、教育現場でのRCTが成立しておらず、得られた結果に剰余変数が影響していた可能性があることを報告している。確かに、現実場面における研究では社会的要因などの影響を除去することが難しく、実

験室実験と異なり、調査者が管理できない領域が多岐に渡るため、剰余変数が多く生まれる。また、これらの要因を統制するためにもランダム化が重要になるが、現実として完全なランダム割り当てを実施することも難しく、十分な標本サイズが得られる保証もない。そして、完全な実験的実験として実施された場合、生態学的妥当性への疑問が生じ、現実世界に適合したエビデンスにはならない可能性もある。

これらと同様の指摘は学習科学という文脈の中でもなされている。学習科学は認知科学研究の知見を学習環境に応用することを目的としており（三宅，2006），実験室ではなく、日常場面における学習を研究するということを中心課題としている（Collins, Joseph, & Bielaczyc, 2004）。Collinsら（2004）は、学習科学が多くの統制不可能な変数が存在する日常の混沌とした場面での学習について研究するものであり、得られた知見が生態学的妥当性の高い環境で機能しなければならないことを指摘している。さらに、彼らは、単独の従属変数への焦点化ではなく、複数の従属変数への同時的アプローチを提唱している。特に、RCTをベースとした実験的デザイン研究の場合、現実的に複数の従属変数を同時に検討することには限界がある。また、独立変数の数も実験デザインに依存するため、その数に制約が生じる。さらに、その分析においても同時に検討できる独立変数の数は限られることになる。

RCTは独立変数の従属変数に対する効果を因果関係の下で検討するための極めて強力なツールであることは間違いない。しかし、必ずしも教育分野でのニーズに適合するものではなく、RCTが適用できるリサーチクエスションも限られることになる。こうした観点から、目的に応じてエビデンスの厳密度を柔軟にして活用するという考え方が教育分野では取られるようになってきている（岩崎，2010）。実際に、イギリスにおいて質的研究や概念的研究の知見をもエビデンスとして含むようになり（惣脇，2012），また、目的や状況に合わせて質的データを含む多様なエビデンスを用いるため、「エビデンス情報に照らした」という

言葉が用いられるようになっている（石井，2015）。つまり、問題の性質や方法論的立場によって最良のエビデンスの在り方が異なるということであり、教条主義的にRCTをベースにしたものだけをエビデンスとして考えるべきではないのである。

## 2. 縦断データとその分析手法

エビデンスを柔軟に定義することは求められるが、どのような研究でもエビデンスとして捉えられてしまう事態は好ましいものではなく、研究としての妥当性を堅持したものをを用いていくという姿勢が求められる。

RCTが重視されるのは、実験デザインとして内的妥当性を強く担保するためである。内的妥当性とは、ある結果が特定の原因によって発生したと確認できる程度を指している（Campbell & Stanley, 1963）。因果関係が示されるには、①結果に対する原因の時間的先行性、②2変数間の相関関係の存在、③2変数間の相関関係が疑似相関でないこと、という3条件が満たされる必要がある（Menard, 1991）。つまり、RCTは、研究の内的妥当性を最大化するように計画された実験デザインの在り方であり、この3条件を確認することができると、その結果は因果関係の検証がなされたエビデンスとして成立するのである。

それでは、RCT以外に内的妥当性を持った研究結果をどのように生み出せばよいだろうか。その方法の1つとして縦断データを用いるということが挙げられる。もちろん縦断データであれば良いというわけではなく、調査デザインやサンプリング、質問紙の構成といった事柄も問われるが、それらを満たした上での縦断データは、相関研究という枠組みの中であっても因果関係を検討するための有力なツールとなる。

縦断データとは、広義において、コホート調査から得られるデータも含まれるが、本稿においては、同一の個体を継続的に観察し記録したデータのことを指し、その意味ではパネルデータと同義である。ただし、厳密な意味でのパネルデータは、

同一の個体を継続的に追跡して得られるデータということであって、観察や測定の時点が同一であることは絶対ではなく、個体間で測定時点が異なるアンバランスド・パネルデータも存在する。その点では、本稿における縦断データとは、個体と時点を固定したバランスド・パネルデータを指している。

相関研究の枠組みの中における調査法は、変数の操作をすることは難しいが、多変量的に変数間の関連性を検討でき、実験室実験に比べて生態学的妥当性を担保しやすいという特徴を持つ。しかし、横断データの場合、因果推定の3条件において、変数間の相関関係の存在を検討することは可能であるが、時間的先行性の条件を満たすことができない。そのため、変数Xと変数Yの間にパスを引くならば、 $X \rightarrow Y$ と $X \leftarrow Y$ 、 $X \leftrightarrow Y$ という形のどれかが成り立つ。また、構造方程式モデリングにおいて、XとYの間にどのようなパスを引いたとしてもモデルの適合度は一般的に違いが生じず、同値モデル (Stelzl, 1986) となり、適合度から因果関係を判断することはできない。つまり、因果関係は理論の見地からの推論に依拠していることになる。それに対して、縦断調査によって得られた縦断データは時間的先行性の条件を満たすことができるため、横断データよりも因果関係の検討に優れている。1時点目で測定した変数Xと2時点目で測定した変数Yの間において、 $X_1 \rightarrow Y_2$ というパスを引くことは、 $X_1 \leftarrow Y_2$ や $X_1 \leftrightarrow Y_2$ というパスを引くことよりも自然な形と考えることができる。そこで、本稿では、縦断データを用いたデータ分析法の中でも教育の文脈でも役立つと考えられる因果関係を検討する分析手法と時系列的変化を検討する分析を概観していく。

## 2.1 因果関係の分析

### 2.1.1 交差遅延効果モデル

パネル調査によって得られた2時点以上での縦断データを用いた交差遅延 (遅れ) 効果モデル (cross-lagged effect model) は因果関係の検討に有効である。1時点目の変数Xを $X_1$ 、変数Yを $Y_1$ 、2時点目の変数Xを $X_2$ 、変数Yを $Y_2$ 、誤差項をe

とした場合、図1にあるモデル図で示すことができる。図1では観測変数を用いて図示したが、潜在変数を含む構造方程式を分析に組み込むことも可能である。構造方程式モデリングの枠組みの中で潜在変数を用いることで相関の希薄化の問題に対処でき、時点が異なる同一変数の誤差相関も仮定できる。また、潜在変数からの観測変数に対する係数に等値制約を課すことによって、同一の潜在変数の時点間における実質の違いを抽出することができる (尾崎, 2003)。

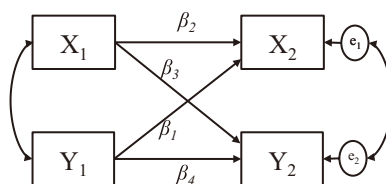


図1 交差遅延効果モデル

2時点目の変数は1時点目の変数によって予測されることになるが、 $X_2$ では $X_1$ が、 $Y_2$ では $Y_1$ が統計的に統制され、その上で、 $X_2$ に対する $Y_1$ の遅延効果、 $Y_2$ に対する $X_1$ の遅延効果が検討されることになる。また、誤差変数間の相関を仮定することで、第3変数の交絡を仮定できる。この事前の値を統制した交差遅延効果から因果関係を検討するという考え方はグレンジャー因果性に対応し (Finkel, 1995)、剰余変数の影響の可能性を排除しきれないことから「グレンジャーの意味で」という言葉がつく因果関係になる。しかし、この分析は変数間の因果関係の方向性を検討する上で有効性を持つものであり、例えば、XがYに影響を与え、その逆ではないということを強く示すことができる (豊田, 2000)。

このように交差遅延効果モデルは因果推論に対して有用なものであり、また、2変数による分析だけでなく、多変量での応用も可能である。そのため、教育分野でも因果関係を踏まえたエビデンスの産出に役立つものであろう。しかし、交差遅延効果モデルは測定時点の間隔が長くなると、真に変数間に因果関係があったとしても、分析からはその因果関係が見出されなくなる場合があり、



適切な分析とはならない。また、時系列的な観点からの因果関係というよりも、同時に生じる因果関係という観点から検討することが適切な場合も存在するだろう。このような場合に対応する分析モデルとして、同時効果モデルがある。

### 2.1.2 同時効果モデル

同時効果モデルは、図2にあるモデル図で示することができる。交差遅延効果モデルと類似した形ではあるが、遅延効果としてのパスが引かれておらず、2時点目のXとYに双方向のパスが引かれている。

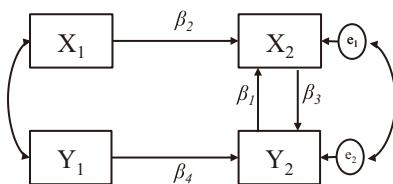


図2 同時効果モデル

また、図2にあるモデル図は誤差変数間に相関が仮定されているフルモデルであり、このモデルの下で $X_2$ と $Y_2$ の相関は、

$$r_{x_2y_2} = \frac{(\beta_1 + \beta_3)}{(1 + \beta_1\beta_3)} + \frac{\beta_2\beta_4r_{x_1y_1}}{(1 + \beta_1\beta_3)} + \frac{\beta_{e_1}\beta_{e_2}r_{e_{12}}}{(1 + \beta_1\beta_3)}$$

という形で構造化され、誤差相関によって第3変数による交絡を分けて考えることができることになる。そして、この形は道具的変数を用いた因果推定モデルと同様の形になっていることが分かる。つまり、同時効果モデルとは、 $X_2$ と $Y_2$ の道具的変数として1時点目で測定した $X_1$ と $Y_1$ を用いたものであり、本質的には道具的変数を用いた因果推定モデルと同様の分析ということになる。この場合、道具的変数が1つの変数にだけ因果的影響を持ち、もう片方には因果的影響を持たないという前提が重要なものになる。この前提が満たされない場合、 $X_2$ と $Y_2$ の間の双方向的にパスを引かなくても相関が説明できるような同値モデルが存在することになり、モデル自体の妥当性が問

われることになる。また、XとYの間に遅延効果が見られるならば、交差遅延効果モデルが成り立つことになる。しかし、同時効果モデルは、因果関係を検討する上で有用なものであることは間違いない。このモデルの枠組みにおいて、適切な道具的変数を用意することができれば、横断的データであっても因果関係の検討が可能になるのである。

## 2.2 時系列的变化の分析

### 2.2.1 潜在曲線分析

縦断データは因果推論に有用であるだけでなく、時間的推移の中の変数の変化を検討することができる。この変化に関する様相を分析するための方法として、潜在曲線分析(latent curve analysis)がある。

潜在曲線分析は、潜在変数を用いた構造方程式モデリングによる確証的因子分析の応用であり、個々人の変化パターンから平均的な変化パターンを捉えることを目的としている(Bollen & Curran, 2006)。潜在曲線分析では、観測変数に対する潜在変数からのパス係数が固定母数となっており、パス係数が解釈の対象ではない。潜在変数の平均や分散、因子間の共分散が重要であり、初期値としての切片と変化としての傾きが解釈の対象となる(豊田, 2007)。つまり、縦断データから変化の形状を表す潜在変数を見出し、それらの母集団における散らばりを潜在変数の平均と分散を用いて表現し、母集団全体における変化の形状を検討するのである(飯塚, 2003)。

まず、縦断データでは、個々人の直線を求められるので、潜在曲線分析において、 $t$ 時点における個人のケースの得点 $y_{it}$ は、切片を $\alpha$ 、傾きを $\beta$ 、誤差を $e$ とすると、

$$y_{it} = \alpha_i + \lambda_i \beta_i + e_{it}$$

となる。この式から分かるように、切片と傾きが個人ごとに異なることを仮定している。そして、ここでの $\lambda$ は、一般的な線形モデルにおいては定数項になり、3時点での縦断データの場合、

$$\begin{aligned}y_{i1} &= \alpha_i + 0\beta_i + e_{i1} \\ y_{i2} &= \alpha_i + 1\beta_i + e_{i2} \\ y_{i3} &= \alpha_i + 2\beta_i + e_{i3}\end{aligned}$$

となる。つまり、 $\alpha$ と $\beta$ を共通因子とする因子分析モデルであり、 $\lambda$ が因子負荷量にあたる。潜在曲線分析は、この共通因子としての潜在変数である切片と傾きの平均や分散を推定することを目的とし、3時点の縦断データの場合、図3にあるモデル図で表現される。

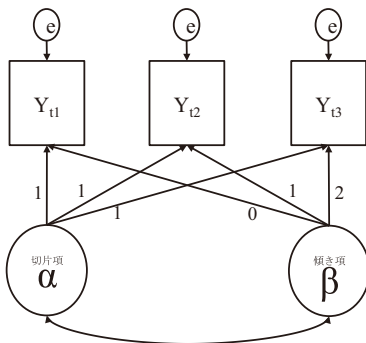


図3 3時点での潜在曲線モデル

潜在曲線分析は、個々人の切片や傾きから変数の平均的な変化の様相を記述できるが、設定した潜在曲線がデータと適合せず、モデル自体の適合度が低いことも考えられる。このような場合、同じ個人からの反復測定による縦断データであり、内容的に同じ観測変数であるため、実質的な意味があるならば、観測変数の誤差間に誤差共分散を導入するという方法がある（飯塚, 2003）。また、別の方法として、高次項を導入し、2次関数的な変化や3次関数的な変化を仮定したモデルで検討するという方法も存在する（Hancock, Kou, & Lawrence, 2001）。例えば、2次関数モデルの場合は、2次項を追加することにより、非線形的な曲線を表現することができる。この場合では、 $t$ 時点における個人のケースの得点は、

$$y_{it} = \alpha_i + \lambda_i \beta_{1i} + \lambda_i^2 \beta_{2i} + e_{it}$$

となる。そして、4時点での縦断データの場合、

$$\begin{aligned}y_{i1} &= \alpha_i + 0\beta_{1i} + 0\beta_{2i} + e_{i1} \\ y_{i2} &= \alpha_i + 1\beta_{1i} + 1\beta_{2i} + e_{i2} \\ y_{i3} &= \alpha_i + 2\beta_{1i} + 4\beta_{2i} + e_{i3} \\ y_{i4} &= \alpha_i + 3\beta_{1i} + 9\beta_{2i} + e_{i4}\end{aligned}$$

となり、図4にあるモデル図で表現される。つまり、2次関数モデルでは、共通因子としての $\alpha$ と $\beta_1$ 、さらに2次項としての $\beta_2$ を推定していくことになる。

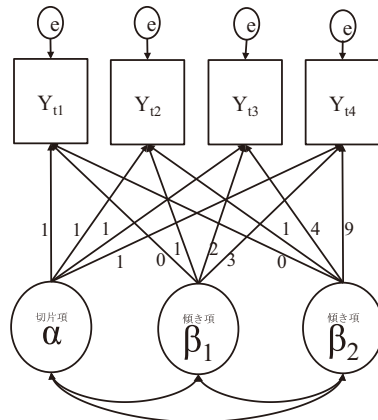


図4 4時点での2次潜在曲線モデル

ただし、2次関数モデルでは、モデルの識別のために4時点以上の縦断データが必要になり、3次関数モデルでは5時点以上の縦断データが必要となる（岡林, 2006）。また、1次関数モデルと異なり、切片や傾きの具体的な意味づけが難しく、誤差変数間の共分散や係数の解釈が難しいという問題がある（豊田, 2000）。さらに別の方法として、傾きからの3時点目以降の観測変数に対するパス係数を自由母数として推定するという方法もある（Duncan, Duncan, Strycker, Li, & Albert, 1999）。この方法は、変化の近似は1次関数を用いながら、曲線的な近似に近い推定が得られる方法であり（豊田, 2007）、通常の1次関数モデルと同様に切片と傾きの解釈も明瞭な形で得られることができる（飯塚, 2003）。

このように潜在曲線分析は、縦断データから線形的・非線形的変化に関する様相、そして、その変化と初期値の関係性を統計的に記述することが

可能である。さらに、潜在曲線分析は、推定された切片と傾きを従属変数や独立変数として扱うことができる。つまり、横断データの場合は、Y自体の値とXの関連を検討することになるが、潜在曲線モデルの枠組みでは、時系列的なYの「変化」とXの関連を検討できるのである。また、潜在曲線モデルに独立変数を組み込むだけでなく、2つの潜在曲線モデルの切片や傾きの間にパスを引くなども可能であり、幅広い応用が可能でもある。

### 2.2.2 グループに基づく曲線

潜在曲線モデルは平均的な切片と傾きを推定ため、異なった潜在曲線を持つ複数の集団が混在すると適合度が低くなる可能性もある(飯塚, 2003)。このような場合、潜在的に異なる集団が既知であれば、多母集団分析での検討を組み込み、平均構造に関する集団差を比較し、集団を分けて検討するということが考えられる。しかし、その集団が未知のものであり、標本がどの程度の数の集団に所属しているか分からない場合もある。このような場合には、潜在曲線混合モデル(latent curve mixture model)が有用である(例えば、Muthén & Shedden, 1999)。潜在曲線混合モデルは、潜在曲線分析と有限混合モデリングを組み合わせた「ような」ものである。

1次関数での潜在曲線混合モデルにおいて、ある集団 $g$ に所属する個人の $t$ 時点における得点 $y_{it}$ は、切片を $\alpha$ 、傾きを $\beta$ 、誤差を $e$ とすると、

$$f(y_{it}) = \alpha_{gi} + \lambda_{it}\beta_{gi} + e_{git}$$

となる。そして、切片と傾きは

$$\begin{aligned}\alpha_{gi} &= \alpha_g + u_{\alpha_{gi}} & u_{\alpha_{gi}} &\sim N(0, \sigma_g^2) \\ \beta_{gi} &= \beta_g + u_{\beta_{gi}} & u_{\beta_{gi}} &\sim N(0, \sigma_g^2)\end{aligned}$$

として説明できる。つまり、ある集団における切片と傾きとその個人差という形で説明され、個人差は正規分布に従うランダム変数である。これらは潜在曲線分析と基本的に変わらないが、集団の違いが含まれている。そして、この潜在曲線モデ

ルの式を有限混合モデリングに当てはめることになる(Muthén, 2002)。有限混合モデリングでは、個人( $i$ )のある測定時期 $t$ における変数である $y_{it}$ が、 $G$ 個の集団からなる有限混合分布に従うとすると、

$$y_{it} = \sum_{g=1}^G p_i^{(g)} \hat{y}_{it}^{(g)}$$

となる(Jones & Nagin, 2007)。 $p_i^{(g)}$ は個人 $i$ が集団 $g$ に所属する確率であり、従属変数の関数である $\hat{y}_{it}^{(g)}$ に潜在曲線モデルが当てはめられる。この時、 $p_i^{(g)}$ は、

$$\sum_{g=1}^G p_i^{(g)} = 1$$

となる。つまり、潜在曲線モデルによって求められる平均的な切片と傾きからの曲線についての形状的な差異を基礎とした潜在的なクラスを求め、個人が潜在的なクラスに所属する確率を検討するということになる。その際、潜在クラス数は未知であるため、情報量基準であるAIC、BIC、CAIC、aBIC(Henson, Reise, & Kim, 2007)、または、EntropyやCLC(Celeux & Soromenho, 1996)を参照し、潜在クラス数を統計的に推定し、その数を分析者自身が決定しなければならない。その後、潜在クラスに含まれた個人の特徴を検証するプロフィール解析によって、クラス自体の特徴を把握できるようになり、個人内変化や見出された変化に関する集団の違いの様相を把握できるのである。

しかし、潜在曲線混合モデルは、集団レベルの固定効果と個人レベルのランダム効果からなるマルチレベル分析であるため、集団と個人の両方に母数推定が求められ、モデルや母数推定が複雑で頑健ではなく、推定における初期値を変えることによって解が変化するということが報告されている(Nagin, 2005)。それらの問題に対して、グループに基づくアプローチであり、NaginとLand(1993; Nagin, 1999, 2005)による混合軌跡モデル(group-based trajectory analysis)が有効である。

混合軌跡モデルでは、ある集団の中で個々人の変数のパターンが等質であるという仮定と軌跡の分割によって集団間での変動性が明らかになるような異なる集団の存在という仮定がある（例えば、Muthén, 2003; Nagin, 1999, 2005）。そのため、混合軌跡モデルにおいて、有限混合モデルでは、誤差項の分散共分散を 0 に固定するという制約が加えられ、集団内の個人差が仮定されない（Muthén, 2004）。つまり、混合軌跡モデルは、個人レベルのランダム効果を考えず、集団レベルのみのモデルであり、各集団の中では全ての個人が同じ係数を持つという制約がある一方、異なる集団では異なる係数を持っていても構わないということになる。そのため、推定する母数の数が潜在曲線混合モデルよりも少なく、頑健な結果が得られるのである。また、混合軌跡モデルでは、打ち切り正規分布やゼロ過剰ポアソン分布、2 値ロジット分布などを用いることができることも特徴として挙げられる（Nagin, 2005）。

集団の数の決定は、基本的に潜在曲線混合モデルと同様であり、情報量基準を参照しながら決定していくことになる。そして、情報量基準を基に、決定された集団ごとの軌跡の関数形を決定する。結果の解釈において重要なのは、切片や傾きからの軌跡の形状、各集団に属する標本のサイズや割合、個人の集団に対する所属確率である。Nagin (2005) は、選択されたモデルが適切に標本における異質性をとらえていることの診断法として、各集団への割り当ての平均事後確率と正分類のオッズ（OCC）を挙げている。平均事後確率は集団の割り当ての精度、または確実性を反映しており、平均事後確率が 1 に近いほど、集団の割り当てはより確実であると考えられる。そして、OCC では、平均事後確率と各集団の推定された割合が用いられ、OCC が大きいほど、正分類のから割り当ての平均確率がランダムな偶然を超えていると考えられ、割り当ての精度は良いと判断できる。

また、混合軌跡モデルにおいて説明変数を導入し、説明変数の値によって各集団への所属確率がどのようになるかを検討することができる。得られる指標としては（多項）ロジット係数とその検

定統計量であり、説明変数の値が 1 大きくなったときの、別の集団への所属対数オッズの平均的増量を示しており、集団の割り当てに対する説明変数の寄与を検討することになる（Muthén, 2004）。このような手続きの下で、混合軌跡モデルは、縦断的データの時系列情報を含んだ形で軌跡を求めることができ、異なった軌跡を持つ異なる集団の存在とその集団への移行を説明する変数について検討することができるのである。

### 3. 教育分野における縦断データ分析の応用

本稿では縦断データの分析手法として、交差遅延効果モデル、同時効果モデル、潜在曲線モデル、潜在曲線混合モデル、混合軌跡モデルを概観した。これらの分析手法を概観することにより、逆説的ではあるが、縦断データを用いることのメリットとして、1) RCT をベースとしたデザインでなくとも、得られた結果から因果推論が可能であること、2) 独立変数の効果としての固定効果と個人差としてのランダム効果を分割して検討できることができ、個人が時間的経過と共にどう変化するかという個人内変化や変化の個人差を検討できること、3) 異なった変化の軌跡とその移行に関わる要因についての検討が可能であること、ということが挙げられる。

RCT との比較として述べるならば、調査によって得られた縦断データによる因果関係の推定力は、剰余変数の影響を完全に取り除くことができないため、RCT をベースとした実験デザインに劣るものではあろう。しかし、調査法の利点として多くのサンプル数を確保した上で複数の変数を測定することができるため、剰余変数を統計的に統制することにより、より高い精度での因果推定も可能である。そのため、縦断データを用いることは、RCT に劣るとも負けない程度に内的妥当性を高めた研究を可能にするだろう。また、変数操作を用いずとも変数間の因果関係を推定できるという点は、RCT をベースとした実験デザインでは倫理的・現実的な観点から研究対象にすることが難しいテーマも検討可能にすることができる。例え



ば、教育介入プログラムの評価研究の場合、参加者に対して一様にプログラムを処遇した上で、その効果を検討することができ、それはプログラムを受けられる学生とそうではない学生に分けられることによる不公平などを生じさせるものではない。つまり、調査法の枠組みの中で縦断データを集めるということは、実践現場においてRCTが持つ方法論上の倫理的問題を軽減することができるのである。

次に、縦断データを用いることによって固定効果とランダム効果を分割して検討できる点は、「エビデンス」という観点において重要であろう。例えば、研究者が検討したいと考える教育介入プログラムの効果が真であるならば、個体間で変わらない要因であるはずであり、個人差であるランダム効果と切り離して推定する必要がある。多時点測定による縦断データの場合、時間的経過による成熟効果なども個人差として切り離すことができ、その結果は、RCTをベースにした実験デザインによって得られたデータからの結果よりも、より真値に近い独立変数の効果の推定を可能にすることができる。また同時に、変化の個人差や異なった変化の軌跡を検討することにより、教育方法や実施したプログラムの改善にもつなげることを可能にするだろう。このように調査法という枠組みの中で得られた縦断データとその分析手法を利用することにより、RCTが持つ弱みをフォローしながらも、その統計的分析結果は、RCTとは異なった側面からエビデンスとして成立しうるものである。

本稿で概観した分析手法は多くの学問領域で活用されている。日本において、交差遅延効果モデルや同時効果モデルは、医療福祉分野（例えば、能村・二木, 2011）や社会学分野（例えば、西岡・星, 2009）などでも多く用いられている。教育分野に関しても、学校生活の諸領域に対する適応とその重要度認知の間の因果関係を検討した研究（岡田, 2012）などが見られ、その活用が大きく期待できる。潜在曲線モデルは主に発達分野で用いられることが多いが、山森（2004）による中学生の英語学習に対する意欲の持続性についての研

究や角谷（2005）による中学生の部活動への積極性と学業コンピテンス、学校満足感の関連性についての研究のような学校教育の文脈の下での使用も見られる。潜在曲線混合モデルと混合軌跡モデルについては、原田（2007）による空き巣の認知件数の研究や中田（2012）による高齢者の健康満足感の推移についての研究があるが、教育の文脈の下での応用研究は日本において少ないのが現状である。海外においては、幼稚園児たちの数技能の発達モデル化の研究（Jordan, Kaplan, Olah, & Locuniak, 2006）や学校関与（School engagement）の軌跡によるドロップアウトの予測についての研究（Janosz, Archambault, Morizot, & Pagani, 2008）など数多くの応用研究が見られる。

日本において潜在曲線混合モデルと混合軌跡モデルの応用研究が少ないことの背景には、これらの分析を実行するための十分な標本サイズの縦断データを得ることが難しいということが関連しているだろう。縦断調査の期間の中で個人を追跡することには金銭的成本、時間的成本、マンパワーが必要となり、往々にして標本の脱落が様々な理由から多く生じる。このような調査コストの高さという点では、RCTと同程度の困難さが生じると考えられ、そのような縦断調査を個人レベルで継続することは難しく、十分な標本を得るまで継続的に調査を行うには組織的な取り組みが求められる。しかしながら、日本では大規模な縦断データを継続的に収集するための組織的体制が海外に比べて整っていないのが現状である。本稿で概観した分析手法以外にも潜在推移分析など多くの縦断データの分析手法が存在し、教育分野でも多くの知見を生み出すことが可能である。教育という実践の場への活用のためには、質の高い縦断データを得るための組織的体制とその支援が必要であろう。

## おわりに

本稿は、エビデンスに基づく教育というテーマに対して、RCTが用いられた研究がエビデンスに成りうる背景を考察し、より柔軟なエビデンスの

産出の在り方の1つとして縦断データの使用とその分析手法を提唱した。教育は必ずしも短期的な因果関係で構成されるものではなく、児童期に経験した事柄が成人期に結果として発現するといった離散的な在り様、または、乳幼児期から青年期へと継続的に関連し合うような連続的な在り様など、長期的な観点から検討しなければ見えない事柄も多々あるだろう。このような長期的な事柄は、RCTのような実験デザインでは捉えることが難しく、縦断的にデータを集め、時系列的情報を踏まえて分析整理することが求められるものである。このことを考えても、時系列デザインによる縦断データからの研究結果は、教育分野におけるエビデンスの産出の在り方の1つとして相応しいものと思われる。

しかしながら、縦断データとその分析手法の使用においても留意すべきことはある。まず、縦断データを用いた研究においても系統的レビューは必要である。研究結果には常に測定誤差と標本誤差が含まれ、結果の一般化可能性が担保されているわけではない。実験室実験とは異なり、多くの変数が交絡した社会的現実の中にある教育分野には様々な状況や標本が存在し、それぞれが異なる文脈の中に埋め込まれている。そのため、得られた研究知見を系統的レビューによって統合整理し、異なる文脈を持つ教育の場に適用できるかを検証することが重要になるのである。

次に、当然のことながら縦断データを用いた分析手法を使えばエビデンスと認められる結果が得られるというわけではない。統計的な分析手法は理論と研究のための道具であり、それを用いることが目的ではなく、縦断データとその分析手法も目的を達成するための選択肢の1つでしかないのである。同様に、RCTも目的達成のための選択肢の1つに過ぎない。その点では、目的に応じてエビデンスの厳密度を柔軟にして活用するという考え方が教育分野では取られるようになってきていることは、正しい方向性に向かっていると言える。

第3点目の留意点として、調査法をベースとした縦断データを用いることは、方法論上の倫理的問題を軽減した上で質の高いエビデンスを産出し

うるであろうが、研究としての倫理的問題を解消するものではないということである。RCTであっても、縦断的調査であっても、インフォームド・コンセントを得ることや実験調査に参加したくない参加者や脱落者への倫理的配慮などは変わるものではない。そして、なにより、研究結果が最終的に学生や教育のためになるとしても、教育実践の場において研究のために学生を用いるという事自体が倫理的問題を内包しており、どのような研究手法を用いたとしても、研究である以上は永遠に解消されないジレンマなのである。

最後に、誰がエビデンスを用いるのかということに留意しなければならない。縦断データとその分析手法は研究者にとっては魅力的なものであろう。しかし、研究者はエビデンスの産出者であって使用者ではなく、使用者とは教育政策の立案者や教育の現場で働く教員である。こういった人たちに対して、エビデンスの産出者である研究者は結果を分かりやすく提示する必要がある。

本稿は、分析手法の統計数学的な部分を極力除き、その概要を述べたが、本質的な理解と適切な使用のためには、統計学の基礎と数学的理解も必要である。それはRCTや質的研究であっても同様であり、方法論の正確な理解が求められる。このような観点から考えると、エビデンスに基づく教育という考え方が普及するのは悪いことではないが、エビデンスという言葉が一人歩きしないように、そのエビデンスの内容を理解し、批判的に検討するための教育体制こそが重要になるのである。

## 引用文献

- Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for over extraction of latent trajectory classes. *Psychological Methods*, 8, 338–363.
- Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. New York: John Wiley.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental design for research*. Hopewell, NJ: Houghton Mifflin Company.

- Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13, 195–212.
- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, 13(1), 15–42.
- Finkel, S. (1995). *Causal analysis with panel data*. Thousand Oaks, CA: Sage Publications.
- Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Albert, A. (1999). *An Introduction to latent variable growth curve modeling: Concepts, issues, and application*. Lawrence Erlbaum Associates.
- Fitz-Gibbon, C. (2000) Education: Realizing the Potential,. In Davies, H.T.O., Nutley, S.M. & Smith, P.C. (Eds.), *What Works?: Evidence-based Policy and Practice in Public Services* (pp. 69–91). The Policy Press.
- 藤澤由和・濱野強・小藪明生 (2007). 地区単位のソーシャル・キャピタルが主観的健康感に及ぼす影響. 厚生学の指標, 54 (2), 18–23.
- Hancock, G. R., Kou, W., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling*, 4, 157–175.
- 原田豊 (2007). 東京23区における空巢の認知件数の軌跡分析--セミパラメトリック混合ポアソンモデルによる検討 犯罪社会学研究 32, 27–44.
- Henson, J. M., Reise, S. P., & Kim, K. H. (2007). Detecting mixtures from structural model differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling*, 14(2), 202–226.
- 飯塚久哲 (2003). 潜在曲線モデルのモデル改善 豊田秀樹 (編著) 共分散構造分析 (疑問編) 朝倉書店, pp. 198–200.
- 石井英真 (2015). 教育実践の論理から「エビデンスに基づく教育を問い直す」: 教育の標準化・市場化の中で 教育学研究, 82 (2), 216–228.
- Janosz, M., Archambault, I., Morizot, J., & Pagani, L. S. (2008). School engagement trajectories and their differential predictive relations to dropout. *Journal of Social Issues*, 64, 21–40.
- Jones, B. L. & Nagin, D. S. (2007). Advances in group-based trajectory modeling and SAS Procedure for estimating them. *Sociological Methods Research*, 35, 254–571.
- Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematical difficulties. *Child Development*, 77 (1), 153–175.
- Long, D. A., & Perkins, D. D. (2007). Community social and place predictors of sense of community: A multilevel and longitudinal analysis. *Journal of Community Psychology*, 35, 563–581.
- Menard, S. (1991). *Longitudinal research*. Newbury Park, CA: Sage Publications.
- 三宅なほみ (2006). 学習科学: 協調的な実践科学と理論構築との互惠関係をめざして 人工知能学会誌, 21 (1), 77–84.
- Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29, 81–117.
- Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling. *Psychological Methods*, 8, 369–377.
- Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D.W. Kaplan (Eds.), *The Sage Handbook of Quantitative Methodology for the Social Sciences* (pp. 345–368). Thousand Oaks, CA: Sage.
- Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55, 463–469.
- Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, 4, 139–157.
- Nagin, D. S. (2005). *Group-based modeling of development*. Harvard University Press.
- Nagin, D. S., & Land, K.C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed poisson model. *Criminology*, 31(3), 327–362.
- 中田知生 (2012). 高齢者における健康満足度の推移と社会階層: 集団軌跡モデルを用いて 医療と社会, 22 (1), 79–89.
- 難波久美子・玉井航太・河合優年・山本初実 (2013, 9). 潜在成長曲線分析を用いた発達経路の探索的検討 (1): KIDS総合発達年齢を用いたモデルの提示 日本心理学会第77回大会発表論文集, 1026.
- 西岡八郎・星教士 (2009). 夫のワークライフバランスが妻の出産意欲に与える影響 人口問題研究, 65 (3), 58–72.
- 能村友紀・二木淑子 (2011). 在宅要介護高齢者における床上動作能力と転倒恐怖感の因果関係の推定作業療法, 30 : 158–166.
- 岡林秀樹 (2006). 発達研究における問題点と縦断データの解析方法 パーソナリティ研究, 15 (1), 76–86.
- 岡田有司 (2012). 学校生活の諸領域に対する適応と重要度認知の因果関係: 交差遅延効果モデルによる検討 パーソナリティ研究, 21 (2), 186–189.
- 尾崎幸謙 (2003). 2時点での因子の違い 豊田秀樹 (編著) 共分散構造分析 (疑問編), 朝倉書店, pp.194–196.
- Sarason, S. B. (1974). *The Psychological sense of community: Prospect for a community psychology*. San Francisco: Jossey-Bass.
- 惣脇宏 (2012). 英国におけるエビデンスに基づく教育政策の展開 国立教育政策研究所 (編) 教育研

- 究とエビデンス, 明石書店, pp.25-42.
- Stelzl, I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research*, 21, 309-331.
- 角谷詩織 (2005). 部活動への取り組みが中学生の学校生活への満足感をどのように高めるか: 学業コンピテンスの影響を考慮した潜在成長曲線モデルから 発達心理学研究, 16 (1), 26-35.
- 豊田秀樹 (2000). 共分散構造分析 (応用編) 朝倉書店
- 豊田秀樹 (2007). 共分散構造分析 Amos編 東京図書
- 津谷喜一郎 (2000). コクラン共同計画とシステムティック・レビュー: EBMにおける位置付け 公衆衛生研究, 49 (4), 313-319.
- 津富宏 (2003). 系統的レビューに基づく社会政策を目指して: キャンベル共同計画の取組み 日本評価研究, 3 (2), 23-39.
- 岩崎久美子 (2010). 教育におけるエビデンスに基づく政策: 新たな展開と課題 日本評価研究, 10 (1), 17-29.
- 山森光陽 (2004). 中学校1年生の4月における英語学習に対する意欲はどこまで持続するのか 教育心理学研究, 52 (1), 71-82.

本研究は科学研究費補助金による「ソーシャル・キャピタルと学校教育の公共性・正統性に関する比較社会学的研究」(課題番号20330176: 研究代表者: 藤田英典) 及び「東日本大震災と教育に関する総合的研究」(課題番号24243073: 研究代表者: 藤田英典) の助成を受けている。