

研究委員会企画チュートリアルセミナー

知っておきたいマルチレベル分析の考え方

ー データの階層構造と文脈情報を活かした分析のために ー

企画者・司会者・

話題提供者 平井 洋子 (首都大学東京)
 企画者・話題提供者 伊藤 亜矢子 (お茶の水女子大学大学院)
 企画者・話題提供者 杉澤 武俊 (大学入試センター*)

*現所属：新潟大学

I. 企画の趣旨

複数の学校や学級から得たデータがあるとき、学校や学級という単位を無視してデータ解析をすると、観測値間の従属性により誤った結果を得るおそれがある。このようなとき、マルチレベル分析(「階層的線形モデル:HLM」などとも呼ばれる)では、学校や学級という上位構造の下に個人のデータがネスト(入れ子)されているという見方をする。そして学校や学級の特徴を表す変数を分析に取り込み、それらを文脈情報としながら、下位(個人)レベルでの分析と上位(学校や学級)レベルでの分析を同時に行う。

本セミナーは、広く教育心理学の研究者を対象とし、マルチレベル分析によってどのようなリサーチ・クエスチョンに答えられるかを紹介することに重点を置いた。そこで、まず先行研究から分析例を簡単に紹介し、どのような構造のデータにこの分析法が適用されるのかを概観した。次に、実際の研究で生じたりリサーチ・クエスチョンと既存のデータ解析法に感じる不満を紹介し、それに対しマルチレベル分析ではどのように考え、モデルを組み立て、解を得るかを事例に即して具体的に示した。そして、研究者と解析者が共同で結果を解釈し、最後に研究者がリサーチ・クエスチョンに即して研究上の意味を確認する、という流れをとった。

II. マルチレベル分析の研究例とデータ構造

文献データベース PsycINFO によると、マルチレベル分析を用いた研究はここ20年ほどの間で飛躍的に増加し、因子分析を用いた研究を100%としたときの割合も、近年では15%から20%に迫る勢いである。その約90%は応用研究であり、研究領域も「学習・学業達成」「対人社会的発達・人格的発達」「教室ダイナミクス・生徒の適応・態度」「カリキュラム・教育プログラムと教授法」など、教育心理学に密接に関連する領域が多い。

マルチレベル分析の主な適用場面として、横断的研究、縦断的研究、メタ分析の3つがある。本セミナーでは、そのうち横断的研究と縦断的研究から1つずつ研究例を紹介した。

横断的研究の例として取り上げたのは Thorpe (2006) である。この研究では、イギリスにおける PISA 2000 の「読解」のデータ(中学3年生7,839人、中学校333校)が用いられた。データの構造は、Figure 1にあるように、学校という上位構造の下に生徒がネストされる、階層的な構造である。

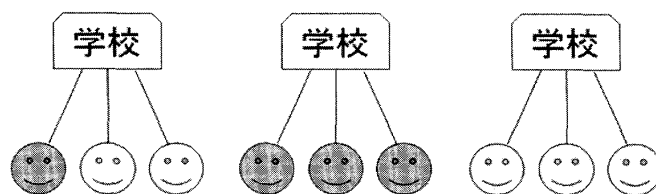


Figure 1 Thorpe (2006) のデータ構造

Thorpe (2006) は、生徒用の質問紙から3因子(家庭の社会的経済的地位(SES)、教育的家庭環境、学校や学級に対する好意的評価)を構成し、その生徒の読解得点を予測するモデルを立てた。分析の結果、同じ学校内における生徒の個人差(下位のレベル)を説明していたのは、家庭のSESや教育的家庭環境などであった。また学校間の得点差(上位のレベル)については、私立校の得点の高さは家庭に恵まれた生徒が集まっているためであり、私立校が特別に良い教育をしているわけではないことがわかった。

縦断的研究の例は、Tsai, Kunter, Lüdtke, Trautwein, & Ryan (2008) である。この研究では、ドイツの中学1年生261人(2校、9学級)に対し、3週間、3科目にわたって毎回の授業が終わるごとに授業を面白く感じたかどうかを尋ねた。データの構造は、Figure 2にあるように、各生徒の下に毎回の測定結果がネストされる階層構造である。分析の結果、同じ生徒・教科でも授業の状況的要因によって面白さが異なったが(下位のレベル)、もともとその教科に興味があった生徒は面白いと感じる度合いが高いことがわかった(上位のレベル)。

マルチレベル分析は、このように階層構造をなしているデータに対して、複数のレベルを同時に解析する。各レベルをどのように設定するか、各レベルを説明する変

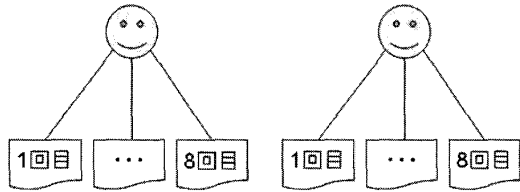


Figure 2 Tsai et al. (2008) のデータ構造

数を何にするか、などによって様々な応用が可能となる。

III. マルチレベル分析の必要性—研究の場面から—

学級の個性を捉える学級風土研究の現場から、分析方法への問題意識と階層構造を生かしたデータ解析法の必要性が述べられた。

「学級風土質問紙」(伊藤・松井, 2001) には、「いじめられている人がいる」と「クラスの人と仲の悪い人がいる」という2項目がある。この2項目間の相関を調べると、高い学級もあれば低い学級もある。相関が低い学級をよく観察すると、いじめられている人と、クラスの人と仲の悪い人が別の人であったりする。一方、相関が高い学級では両者が関連していたりする(Figure 3)。どちらも学級の個性が反映された結果であり、それを無視して学級の区別を取り払った分析をしても意味がない。

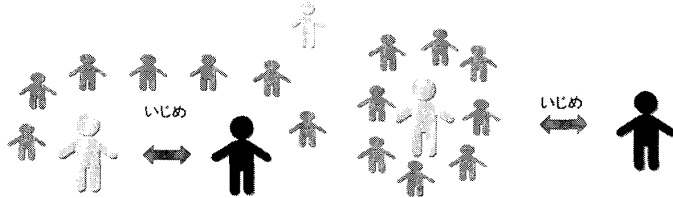


Figure 3 2項目の相関が低い学級(左)と高い学級(右)

また、各生徒の学校への意欲の高さは、学級のあり方や個人の学級に対する認知によって左右されると考えられる。質問紙への回答が影響されるだけではない。同じ回答であっても、その生徒がどのような学級にいるかによって、数値の意味が異なるかもしれない。例えば、「自分の学級における生徒間の親しさ」という項目の学級平均が+1点のクラスで-1点と答えた生徒と、学級平均が-0.5点のクラスで-1点と答えた生徒とでは、値の持つ意味が異なるであろう。ここでも、学級という単位を無視した分析では、実態をつかみ損なうおそれがある。

この学級風土データでは、学級という上位のレベルと生徒という下位のレベルが相互に影響しあっている。どちらかだけのレベルにデータを編成しても、実態を正確に理解することはできない。手元には、中学生2,465人(85学級)分のデータがある。このデータにマルチレベル分析を適用したら、どのような結果が得られるだろうか。リサーチ・クエスチョンは、「『学級への関与』や『生徒間

の親しさ』が『学校への意欲』に影響するか」である。

IV. マルチレベル分析を適用する—解析の側から—

データにネスト構造があるとき、個々の集団で分析した結果と、それらを混在させたときの結果とが異なることがある。統計的にも、同じ集団にネストされた観測値間は互いに似ている確率が高く、母集団からの無作為抽出サンプルといえなくなるため、第1種の誤りの確率や検定力に影響が出る。

学級風土データ全体で「生徒間の親しさ」から「学校への意欲」を予測する単回帰分析を行うと、

$$\hat{y} = 1.28 + 0.68x$$

となるが、3つの学級を選んで学級ごとに単回帰分析を行うと、

$$\text{学級 1 } \hat{y} = 2.53 + 0.20x$$

$$\text{学級 2 } \hat{y} = -0.77 + 1.18x$$

$$\text{学級 3 } \hat{y} = 4.78 - 0.36x$$

となった。つまり学級1は生徒間の親しが増すと学校への意欲も増すが、その変化率は0.2と弱い。学級2は、生徒間が親しくないときには学校への意欲がかなり低いが、生徒間が親しくなると学校への意欲も大きく上昇する。一方学級3は、生徒間が親しくないとき学校への意欲が高く、親しが増すほど学校への意欲も低くなる。これらの状況は、全体での結果($\hat{y} = 1.28 + 0.68x$)に現れた状況とは異なるものである。

この例のように、下位のレベル(ここでは各学級内で分析)における切片や傾きには、その集団の特徴が反映される。マルチレベル分析では、集団ごとの切片や傾きを従属変数とし、集団の特性を記述する変数からその違いを説明しようとする。これを一般的な式で表せば、

下位のレベル(レベル1): 集団内での関係を記述

$$(\text{従属変数}) = \beta_0 + \beta_1 \times (\text{独立変数}) + \text{残差 } r$$

上位のレベル(レベル2): 切片や傾きの集団間の違いを記述

$$\beta_0 = \gamma_{00} + \gamma_{01} \times (\text{集団の特性}) + \text{残差 } u_0$$

$$\beta_1 = \gamma_{10} + \gamma_{11} \times (\text{集団の特性}) + \text{残差 } u_1$$

となる。学級風土データでは、まず以下のモデルを検討する。

分析モデル(1)

レベル1:

$$(\text{意欲}) = \beta_0 + \beta_1 \times (\text{親しさ}) + \text{残差 } r$$

レベル2:

$$\beta_0 = \gamma_{00} + \text{残差 } u_0$$

$$\beta_1 = \gamma_{10} + \text{残差 } u_1$$

ここで切片 β_0 が「学校への意欲」の学級平均を表すようにするため、「親しさ」を学級ごとに学級平均からの偏差

に変換した (group-mean centering という)。傾き β_1 は「親しさ」が「学校への意欲」に及ぼす影響の大きさである。 γ_{00} は切片 β_0 の全学級を通した平均的な値、 γ_{10} は傾き β_1 の全学級を通した平均的な値である。残差 u_0 と u_1 の標準偏差は、それぞれ切片 β_0 や傾き β_1 に学級ごとの違いがどの程度あるかを表す。データから $\gamma_{00}=3.56$ となり、「学校への意欲」の学級平均 (β_0)は平均的に3.56点というといえる。残差 u_0 の標準偏差は0.28なので、学級平均の値は0.28程度の標準偏差をもって学級間で変動しているといえる。一方、 $\gamma_{10}=0.70$ であったので、全学級の平均でいうと、「親しさ」の得点が1点高いと「学校への意欲」が0.70点高くなる。残差 u_1 の標準偏差は0.10であったので、この0.70という傾きは0.10程度の標準偏差をもって学級間で変動しているといえる。

この分析で「親しさ」と「学校への意欲」の関係に学級差があることが確認できたので、つぎにその差を学級の特性を表す変数から説明することを試みる。

分析モデル(2)

レベル1:

$$(\text{意欲}) = \beta_0 + \beta_1 \times (\text{親しさ}) + \text{残差 } r \quad (\text{a})$$

レベル2:

$$\beta_0 = \gamma_{00} + \gamma_{01} \times (\text{関与の学級平均}) + \text{残差 } u_0 \quad (\text{b})$$

$$\beta_1 = \gamma_{10} + \gamma_{11} \times (\text{親しさの学級平均}) + \text{残差 } u_1 \quad (\text{c})$$

このモデルは、「学校への意欲」の学級平均 (β_0)における学級差を関与が高い学級かどうかによって説明し、「親しさ」が「学校への意欲」に及ぼす影響 (β_1)における学級差を、親しい学級かどうかによって説明しようというものである。

分析の結果、 $\gamma_{01} = 0.46$ となり、「関与の学級平均」が1点高い学級は、「学校への意欲」の学級平均 (β_0)が0.46点高いことがわかった。また $\gamma_{11} = 0.18$ となったことから、「親しさの学級平均」が1点高い学級では、「親しさ」が「学校への意欲」に及ぼす影響 (β_1)も0.18だけ強くなることがわかる。ただし残差 u_0 の標準偏差が0.21であったことから、「学校への意欲」の学級平均 (β_0)の学級差を「関与」だけで説明することはできず、まだ0.21程度の説明されない変動 (標準偏差)が残っていることがわかる。同様に残差 u_1 の標準偏差も0.09あり、「学校への意欲」に対する「親しさ」の影響の大きさ (β_1)も、「親しさの学級平均」だけでは説明しきれなかったといえる。

この結果を、研究場面に即して解釈してみよう。

- ・「関与」の高い学級では、「学校への意欲」が総じて高くなる (式(b))。このことは、行事などへの取り組みのよい学級づくりが有効であることを示唆する。
- ・しかし、そこには個人の学級認知も関わっており、「親しさ」の高い学級と感じている子は、「学校への

意欲」が高い (式(a))。そこで、仲の良い学級づくりが望まれる。

- ・「親しさ」の高い学級では、「親しさ」認知の影響がより強く出る (式(c))。例えば「親しさ」の高い学級で「親しさ」認知の低い子は、「意欲」がより低くなる。このことから、「親しさ」が高い学級になるほど、疎外感を感じる子が出ないように気を配る必要があるといえる。

V. マルチレベル分析の注意点

マルチレベル分析では、独立変数をそのまま入れるのではなく、そこから平均値を引いた偏差に変換することが多い。これを変数のセンタリングと呼ぶ。センタリングには、group-mean (集団ごとの平均) centering と grand-mean (全体の平均) centering とがある。センタリングは多重共線性の対策になるとされているが、センタリングの方法やセンタリングの有無によって切片の値の解釈が異なってくるので、よく考えて選びたい。

またレベルの設定であるが、データがどのような階層構造を持つのかを見極めて、研究デザインや分析モデルを決める必要がある。独立変数はレベルごとに必要な変数を測定しておく。ただし、独立変数を入れすぎると計算や解釈の際に問題が生じやすいので、注意したい。

サンプルサイズに関しては、全体の N が多いことも重要だが、上位レベルのユニット (例:学級) 数を増やすことも重要である。

VI. テキストとソフトウェア

統計に詳しくない人のための入門書としては、Bickel (2007), Luke (2004) などが挙げられる。Bickel (2007) は学習者にとってわかりやすいように工夫されており、Luke (2004) は具体的な例をもとにして基本的な事項がわかりやすくまとめられた小冊子である。定評のあるテキストとしては、Raudenbush & Bryk (2002) が挙げられる。マルチレベル分析を勉強するなら一度は読みたい一冊であり、最も普及している解析ソフト HLM の並行テキストでもある。

和書の解説書がほとんどない中、Kreft & de Leeuw (1998 小野寺編訳 2006) は貴重な和書といえる。この本の最終章には、和書だけに追加された部分として解析ソフトの解説がある。主なソフトの操作のしかたや出力の読取り方、関連 web サイトなどが丁寧に紹介されている。

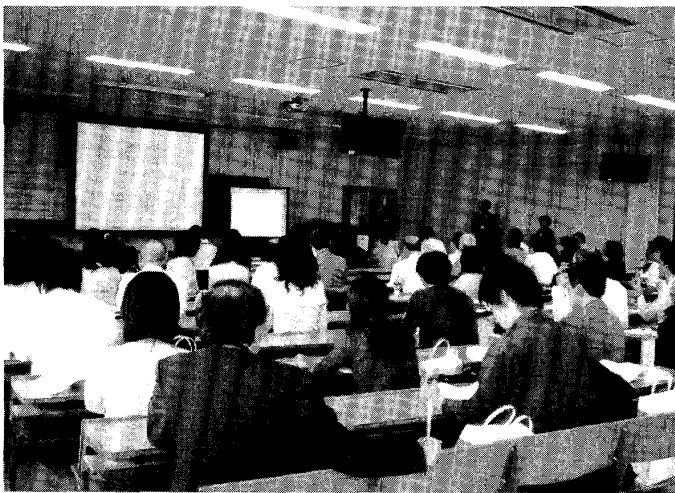
テキスト以外では、マルチレベル分析を紹介した論文も参考になる。宮崎 (2007) は、この分析法の応用分野や歴史の紹介のほか、NELS データを用いて横断的研究と縦断的研究の2通りの適用をわかりやすく丁寧に解説し

ている。その他, Nezlek & Zyzniewski(1998)や Ciarleglio & Makuch (2007) は横断的研究への適用についての紹介論文であり, Maas, Hox, & Lensvelt-Mulders (2004) は縦断的研究への適用についての紹介論文である。

データ解析用のソフトウェアは, 汎用ソフトと専用ソフトの2つの系統に分かれる。汎用ソフトには SPSS の混合モデルや R があるが, 一般には専用ソフトが使われるようである。専用ソフトの代表的なものは, HLM (Scientific Software International) と MLwiN (Centre for Multilevel Modeling) である。先行研究では HLM が用いられる場合が多く, 論文におけるモデル式の記述や結果の記載のしかたも HLM の出力にならっていることが多い。また入門用のテキストも HLM での式や変数の記法に即して記述されていることがあるので, これから専用ソフトを使い始めるなら, HLM がよいかもしれない。

VII. 当日の様子

当日は, ソフトウェア HLM を用いてデータ解析の実演も行われた。またスクリーンを2台使って, データの構造やモデル式, そして解析結果を並べて示した。そのほか, データからモデルを組み立てる練習問題も出題され, 参加者は熱心に取り組んでいた。会場はほぼ満員になり, 本学会員の知識欲や向上心の高さが窺えた。



引用文献

- Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!* New York : Guilford Press.
- Ciarleglio, M. M., & Makuch, R. W. (2007). Hierarchical linear modeling : An overview. *Child Abuse & Neglect*, **31**, 91-98.
- 伊藤重矢子・松井 仁 (2001). 学級風土質問紙の作成教育心理学研究, **49**, 449-457.
- Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. London : Sage. (クレフト, I・デ リュー, J. 小野寺孝義 (編訳) (2006). 基礎から学ぶマルチレベルモデル—入り組んだ文脈から新たな理論を創出するための統計手法— ナカニシヤ出版)
- Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks : Sage.
- Maas, C. J. M., Hox, J. J., & Lensvelt-Mulders, G. J. L. M. (2004). Longitudinal meta-analysis. *Quality and Quantity*, **38**, 381-389.
- 宮崎康夫 (2007). 階層線形モデル (HLM) の教育研究への応用と分析結果の教育政策への利用の観点 日本テスト学会誌, **3**, 123-146.
- Nezlek, J. B., & Zyzniewski, L. E. (1998). Using hierarchical linear modeling to analyze grouped data. *Group Dynamics : Theory, Research, and Practice*, **2**, 313-320.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models : Applications and data analysis methods* (2nd ed.). Thousand Oaks : Sage.
- Thorpe, G. (2006). Multilevel analysis of PISA 2000 reading results for the United Kingdom using pupil scale variables. *School Effectiveness and School Improvement*, **17**, 33-62.
- Tsai, Y., Kunter, M., Lüdtke, O., Trautwein, U., & Ryan, R. M. (2008). What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology*, **100**, 460-472.