

マルチレベル・モデリングによる NFRJ データの分析方法

—ダイアド集積型家族調査の有効活用—

保田 時男

(大阪商業大学総合経営学部)

【要旨】

本研究の目的は、NFRJ データをマルチレベル・モデリングで分析する際の方法論的な留意点を、具体的に考察することである。NFRJ データは、回答者を中心とした 11 のダイアド関係の情報を収集したダイアド集積型の家族調査なので、それぞれのダイアドを第 1 水準、回答者を第 2 水準とするマルチレベル分析が有効である。

ダイアドデータ分析に関する方法論的な研究を土台として、NFRJ のマルチレベル・モデリングの方法について検討した結果、以下のことが明らかになった。第 1 に、1 人の回答者に属するダイアド数のサイズが小さいことから、ランダムな傾きのモデルを安定的に推定することは困難で、ランダム切片モデルに分析を集中させるべきである。第 2 に、ダイアド間に負の相関が発生する可能性があるため、ICC の解釈に注意が必要である。第 3 に、第 1 水準の独立変数の効果には、第 2 水準の効果が強く混入する可能性があるため、解釈が難しい。最後に、世代間援助行動の分析事例によって、NFRJ のマルチレベル・モデリングから鮮明で有効な分析結果が得られることを例証した。

キーワード： マルチレベル分析、ダイアドデータ、家族調査、方法論

1. 目的

NFRJ (National Family Research of Japan; 全国家族調査) は、回答者の視点から複数のダイアド関係(dyadic relation)を並列的に捉えるデータ構造を持っている。回答者からみると、配偶者、4 人の親 (本人の父母、配偶者の父母)、3 人までの子ども、3 人までのきょうだいからなる最大 11 人の家族・親族について、属性と関係性を個別に測定している (図 1)。しかも、配偶者以外のダイアドについては、親も、子も、きょうだいも、調査項目の共通性が高く、異なる種類の家族・親族関係についても同じ側面から捉えられている。このように、回答者を中心にして複数のダイアド関係の情報を積み重ねることでその者を取り巻く家族・親族の状況を把握しようとする調査形式を〈ダイアド集積型〉の家族調査と呼ぶことにしよう。

ダイアド集積型の家族調査である NFRJ は、1 人の回答者に付随する複数のダイアドをそれぞれ個別の観察ケースとみなせば、1 票の調査票が最大 11 のダイアドケースを保有していると考えることができる。たとえば NFRJ08 データ (ver. 4.0) の場合、5,203 人の回答者

があるので、それぞれから 11 のダイアドケースを導くと最大で $5,203 \times 11 = 57,233$ ケースのデータが構成できる。実際には、該当のダイアド関係が存在しない場合があるので、38,818 ケースのデータとなっている¹。

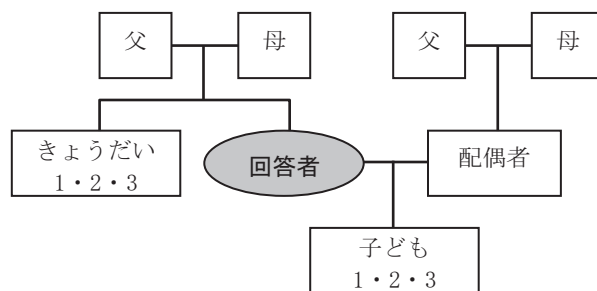


図1 NFRJにおける回答者から見たダイアド関係

このような構造のデータは、マルチレベル分析 (multilevel analysis) の典型的な適用対象となる。マルチレベル分析は、理論的あるいは統計的に複数水準の階層構造を持つようなデータをそのまま複数水準にモデル化 (multilevel modeling) する分析技法である。ここでは、1人の回答者に複数のダイアドケースが包含されることになるので、各ダイアドを第1水準 (level-1)、各回答者を第2水準 (level-2) とする2層のマルチレベルデータを構成していることになる。それぞれ下位水準 (lower level)、上位水準 (higher level) と呼ぶこともある。1人の回答者には最大11のダイアドケースが含まれるが、実際には、ダイアドの種類 (親、きょうだいなど) を限定する研究関心が多いと考えられるので、たとえば、子どもとの関係に限定して最大3個の子ダイアドを第1水準、親 (回答者) を第2水準として、分析を行うような状況が考えられる。

NFRJは、このように典型的にマルチレベル分析の対象となるデータを構成しているが、そもそもそのような意図で調査設計がなされたわけではない。個人を単位とする家族の把握を意図して汎用家族調査を設計した (渡辺 2001) 結果として、このような特異なデータが収集された。そのため、NFRJデータは間違いなくマルチレベル分析が有効なデータ構造を示しているにもかかわらず、その利点や具体的な適用方法、あるいはその限界は明確に提示、議論されていない。

昨今の計量的な家族研究の発展において、マルチレベル分析が果たした役割は甚大であり (Teachman and Crowder 2002)、その勢いは年々強くなっている。NFRJデータがマルチレ

¹ 配偶者 4,062、実親 10,406、義親 6,336、子 8,912、きょうだい 9,102 からなる。ただし、このケース数は単純にデータが記録されている数であり、調査設計上の都合が影響している。たとえば、子やきょうだいは生存している者のみ (最大3人) にデータが限られるのに対して、親は死亡している者も含む。また、配偶者と離死別した回答者や、63歳以上の高齢者には、配偶者の親の情報を尋ねていない。

ベル分析に非常に適したデータ構造を持つにもかかわらず、その適用方法が明らかにされていないために十分に活用されていないとするならば、それは非常に残念なことである。筆者の知るところで、NFRJ データに（上で示したような）マルチレベル分析を適用した分析例は、海外の研究者によるきょうだい学歴の分析事例（Lee 2009）のみである。

本稿の目的は、ダイアド集積型の家族調査である NFRJ データにマルチレベル分析を適用する際の方法論的な留意点を、具体的に考察することである。NFRJ データはいくつかの点で特異な特徴を持っており、一般のマルチレベル分析ではあまり留意する必要のない重要な注意点をもつ。逆に、一般的なマルチレベル分析では重視されるが、NFRJ の分析に限っては留意する必要がない、という点もある。幅広い適用対象を持つマルチレベル分析について、NFRJ への適用に絞った議論を行うことは、その意味で有用である。

第 2 節では、マルチレベル分析の考え方を、とくにダイアドデータ分析 (dyadic data analysis) からのアプローチに力点を置いて整理する。第 3 節では、重要な分析概念に沿って、NFRJ データにおけるマルチレベル分析の留意点を議論する。第 4 節では、ここまでの議論に従って具体的な分析事例を示し、第 5 節で結論をまとめる。

2. マルチレベル分析の概要

マルチレベル分析は、複数の水準のサンプリングが階層的に折り重なることでデータが構成されていると考えるデータ分析法である。3 水準以上のモデルも想定することができるが、ここでは 2 水準のモデルに限定して、まずマルチレベル分析の考え方を簡単に確認しよう。詳しい解説書は多数あるが、入門書としては Luke (2004) や Snijders and Bosker (1999) がわかりやすい。この節で示すマルチレベル分析の概要も、主にこれらに依拠している。

たとえば、中学生の調査をするために、まず学校を単位として、いくつかの中学校をサンプリングし、抽出された学校の中から生徒個人をサンプリングしたとする。サンプリングが無作為に行われたとしても、このデータをそのまま通常の回帰分析 (OLS) 等で分析することには、統計的に問題がある。明らかに i.i.d. の前提に違反しており、生徒個人の抽出が非独立 (nonindependence) に行われているからである。つまり、学校のサンプリングによって、同じ学校に通う生徒は同時に選ばれやすく、違う学校に通う生徒は同時に選ばれにくくなっている。このことを考慮せずに、通常の統計的技法を適用すると、適切な確率計算が成り立たないので、統計的な有意性が過剰に評価される、等のバイアスがかかってしまう。

一方で、この種のデータは理論的な関心を刺激する豊富な情報を含んでいる。ここでの例でいえば、生徒の成績が学校ごとにどう異なるのかを知るためには、同じ学校に通う生徒を複数調べ、成績の共通性 (平均) を学校間で比較しなければならない。逆に、学校に依存しない生徒の個人的な特徴と成績の関係を調べる場合にも、同じ学校に通う生徒同士を直接的に比較する方が有効なので、やはり上記のような 2 水準でのサンプリングに意味

がある。

マルチレベル分析は、このように理論的に必要となる非独立なデータを、統計的に適切に扱うことを可能にする分析技法である。このような発想自体は古くからあるものだが、旧来の分析方法に比してマルチレベル分析には多くの優位な点があり（複数の分析結果を組み合わせる必要がないこと、欠損を含むデータの扱いが容易なこと、等）、マルチレベル分析のための専用ソフト（HLM、MLwin など）や汎用統計ソフトへの組み込み（SPSS や STATA など）が進んだこともあり、その地位はゆるぎないものになっている。

あるデータに対してマルチレベル分析を適用することが有効であることを示すもっとも根本的な基準は、ケース間の非独立性である。非独立でないのならば、マルチレベルを想定する必要はない。非独立性の強さは、統計的には **intraclass correlation coefficient (ICC)** によって測定される。ICC は同じグループに属するケース間の相関であり、ここでの例では、同じ学校に通う生徒の成績の相関係数である。通常、同じグループに属するケースは類似性を示すので、ICC は正の値を示す。また、ICC はその変数の全分散のうち、第 2 水準のグループの違い（つまり、どの学校か）によって説明できる分散の割合を示す。ICC が 0 近ければ、同じグループに属していても相関がなく、したがって非独立とみなす必要がないので、マルチレベル分析を行う意味はない。

非独立性が、統計的、あるいは理論的に確認されたならば、第 2 水準のグループごとに異なるサンプルとして扱うことが適切ということになる。たとえば、生徒個人の成績を勉強時間や家庭背景などの要因で説明する回帰分析を考える場合、学校ごとに異なる回帰式を推定することが妥当ということになる。とはいえ、常識的に考えれば、学校間でも回帰式はある程度類似していると考えられる。マルチレベル分析では、学校ごとの回帰式の違いがどこにどの程度存在し、その違いは学校の属性でどのように説明できるのか、といったことを幅広くモデル化することができる。

多様なモデルが考えられるが、一般に、まず 3 つのモデル群の違いが強調される。**null model (unconstrained model; empty model)** は、学校間での回帰式の違いはまったくない、と考えるモデルである。非独立性が示されている時点で **null model** は妥当ではないわけであるが、他のモデルの参照点として推定されることがある。次に、回帰式の切片だけがグループごとに異なる **random intercepts models** が検討される。つまり、学校が異なっても生徒の勉強時間や家庭背景の影響の仕方自体は変わらないが、ベースとなる平均的な成績が異なるというモデルである。最後に、より複雑なモデルとして、回帰式の切片に加えて傾きも異なるという、**random intercepts and slopes models** が検討される。つまり、学校が異なれば、生徒個人の勉強時間等の影響の仕方も変わってくるというモデルである。複数の独立変数があれば、それぞれの傾きについて異質性の有無が検討される。

さらに、切片や傾きが学校によって異なると考えた場合に、それぞれの切片や傾きの違いを各学校の属性で説明するモデルが考えられる。たとえば、進学校では平均的に成績がよい（切片が大きい）、田舎の学校では勉強時間の効果が出やすい（勉強時間の傾きが大き

い) といったことをモデル化することができる。非独立性を無視した通常の分析でも、独立変数に学校の属性を投入すれば、同様の分析ができるが、統計的に不適切なため、これらの第 2 水準の属性の効果を過大に評価してしまうことになる。マルチレベル分析では、このように理論的に検証が望まれる多様なモデルを、統計的に適切で、かつ比較的容易に推定することができる。

3. マルチレベル分析の視点からの NFRJ データの位置づけ

3.1 ダイアドデータ分析

さて、ダイアド集積型調査としての NFRJ データは、回答者が無作為にサンプリングされた上で、その回答者が関係する複数のダイアドのデータが収集されている。形式上、個々のダイアドが第 1 水準（下位水準）、回答者が第 2 水準（上位水準）のマルチレベルデータである。第 1 水準でランダムサンプリングが行われているわけではないが、前節で示したとおり、重要なことはグループ内の非独立性である。同じ回答者に関係するダイアド、すなわち同じ家族・親族に属するダイアド間に非独立性が存在することは明らかであり、マルチレベル分析が必要とされるデータ形態である。

マルチレベル分析の応用範囲は非常に広いので、その一般的な分析方法から直接に NFRJ データの分析方法を考察することは、アプローチが遠すぎる。そこで、ここでは Kenny et al. (2006) によるダイアドデータ分析 (dyadic data analysis) を土台にして議論を展開したい²。ダイアドデータ分析は、夫婦関係、友人関係、医者・患者関係といったダイアド関係の分析方法を総合的に整理したもので、個人単位を基本に収集されている統計データからいかにして適切にダイアド関係についての知見を取り出すかを精緻に考察している。たとえば、夫婦のそれぞれから夫婦関係についての意見を収集すると、2 つのデータは非独立である。同様に、同じ行為者が取り持つ複数の友人関係は非独立であり、同じ医者から見た複数の患者との関係も非独立である。この非独立性に適切に対処しなければ、ダイアド関係について有効な知見を得ることができない。

簡単にその中心概念を確認しよう。まず、ダイアドデータ分析のデザインは大きく 3 つに分類されている。standard design はすべての個人が 1 つのダイアドだけに属しているようなデザインである。たとえば、夫婦関係データの分析などがこれにあたる。次に、social relations model (SRM) design は各人が複数のパートナーを持ち、その人々もまた複数のパートナーを持つようなデザインで、友人関係のネットワークのような場合がこれにあたる。3 つ目の one-with-many design は 1 人の中心者が複数のパートナーを持ち、その人々は他のパートナーを持たないようなデザインで、1 人の医者が複数の患者との間に取り持つ医者・患者関係などがこれにあたる。ダイアド集積型データとしての NFRJ は典型的に

² Kenny et al. (2006) は、ダイアドデータの分析方法をマルチレベル分析に限らずに非常に多岐にわたって検討しているが、マルチレベル分析に絞った簡略な論考としては、Kenny and Kashy (2011) がある。

one-with-many design にあたる。

次に、それぞれのダイアド関係のデータが一方的 (one-sided) なのか相互的 (reciprocal) なのかが区別される。夫婦関係の調査などで、夫と妻の双方がその関係を同様の調査票で評価していれば相互的なデータといえる。one-with-many design では通常、データは一方的なものであり、NFRJ も回答者視点からの一方的データである。相互的なデータに比べると情報が少ないが、分析手続きの複雑さが小さい。

また、ダイアド間の分別可能性 (distinguishability) も、分析方法を考える際の大きな基準になる。たとえば、友人関係のデータの場合、友人 A との関係と友人 B との関係は分別されない。データ上で 2 つを入れ替えても差し支えないという意味である。一方で、親子関係の研究で同じ子どもに父との関係、母との関係をそれぞれ尋ねている場合、これらは入れ替えることのできない (分別される) データである。NFRJ のダイアド集積型データが分別可能かどうかは、分析目的によって変わってくる。たとえば、複数の子どもとのダイアドについて、出生順位でとくに区別することなく親子関係の分析をするのであれば、分別不能 (indistinguishable) なデータとしての分析になるが、長子と次子の区別を重視して、両者の関連を調べるのであれば分別可能 (distinguishable) なデータとしての分析になる。このように分析内容に依存するものの、ここでは分別不能なデータとして NFRJ の分析方法を考察する。それは、分別可能なデータは、比較的扱いが容易で、分別される各ダイアドを別変数とした構造方程式モデル (SEM) が適用できるからである。Kenny et al. (2006) では、基本的に分別可能な場合は SEM、分別不能な場合にはマルチレベル分析の適用が勧められている。ここでは、マルチレベル分析を適用する分別不能データの扱いに焦点を絞る³。

以上のように、ダイアドデータ分析の視点からダイアド集積型データとしての NFRJ を整理すると、次のようになる。one-with-many design からなる one-sided のダイアドデータであり、分別不能としてマルチレベル分析を適用することが向いている。

さて、ダイアドデータ分析では、多くの点で一般的なマルチレベル分析の考え方との共通性が示されるが、概念的に特筆すべき点が 2 つある。1 つ目は、非独立性がどのような理由で発生しているのか、その源泉へのこだわりである。同じグループ内の複数のダイアド (あるいは、同じダイアド内の 2 人の個人) には、当然、非独立性が予想されるが、ダイアドデータ分析ではその源泉が 4 種類あると考える。第 1 の源泉は、compositional effect であり、これはそもそもダイアドを構成する前から似た者同士が自然にダイアドを構成することが多いという理由である。たとえば、夫婦は夫婦になる前から出身階層が似ている。同じ医者を頼る患者は、似たような性向を持っている。非独立性の第 2 の源泉は、partner effect である。これはダイアドを構成したことによって、直接的に一方が他方の行動や意識

³ ここで、NFRJ データを分別不能とみなすからといって、たとえば分析の独立変数に長子ダミーは投入しない、といったことを意味するわけではない。分別不能なデータとして扱うという意味は、長子ダイアドと次子ダイアドの関係性に注目するわけではないから、これらがワンセットに揃っている必要はない、ということである。逆に、分別可能なデータとして扱うという場合には、注目する分別可能なダイアド群が揃ったデータを扱うことになる。

に影響を与えるということである。たとえば、配偶者の家事頻度が夫婦関係満足度に影響するというような場合である。第3の源泉は、**mutual influence** であり、たとえば夫婦関係満足度を夫婦双方が評価する場合に、相手が満足しているから自分も満足というように相互に影響を与え合うことによる非独立性である⁴。第4の源泉は、ダイアド間の直接的な影響ではなく、同じ環境にさらされていることによる **common fate** である。たとえば、親が共通の複数の親子ダイアドは、同じ祖父母を共有していることになる。同じ地域に住む住民は、相互に交流がなくとも、同じ住環境にさらされている。

一般的なマルチレベル分析では、非独立性の源泉として、**compositional effect** と **common fate** を想定していることが多い。ところが、非常に限定的な範囲でのやりとりを分析するダイアドデータ分析では、**partner effect** や **mutual influence** を考慮しなければならない。ここで注意すべきことは、**partner effect** や **mutual influence** が正の相関だけでなく、負の相関を導くことがある、ということである。たとえば、特定の子どもと親密にしている親は他の子どもと疎遠になる、とか、家事の分担では一方の分担が大きいと他方の分担が少なくなる、といった場合である。通常マルチレベル分析では、**compositional effect** と **common fate** による正の相関を想定しているが、ダイアドデータ分析では負の相関が強い場合もあり、ICC がマイナスの値を示すことも想定しなければならない。

ダイアドデータ分析におけるもう1つの注目すべき概念は、独立変数の種類をその性質から **between variables** と、**within variables**、**mixed variables** の3種類に区別している、という点である。**between variables** とは、第2水準のグループが同じであれば値が同じになる変数のことで、NFRJのマルチレベル・モデリングでいえば、複数のダイアドの中心となる回答者の属性は、**between variables** である。次に、**within variables** は純粋に第1水準の差異だけを示す変数のことである。「純粋に」とは、第2水準のグループの違いが全く影響しないという意味で、たとえば、NFRJで複数の子どもそれぞれとの接触頻度を表す変数は、**within variables** ではない。なぜならば、接触頻度は子どもによって異なる(第1水準の差異を示す)が、そこには、そもそもその回答者が全体的にどの程度子どもと接触しようとするか、という第2水準のグループによる違いも混ざっているからである。このような変数は **mixed variables** に分類される。一方、複数の子どもとの接触回数全体のうち、第1子との接触が何%で、第2子との接触が何%で、……といった変数であれば、第2水準の影響は受けないので、**within variables** ということになる。この例からわかるように、NFRJのマルチレベル・モデリングにおいて、第1水準で測定されている変数(各ダイアドの属性や関係性)は、ほぼすべてが **within variables** ではなく、**mixed variables** である。

第1水準の変数が **mixed variables** であることは重要な意味を持つ。それはつまり、何らかのマルチレベル分析によって第1水準の独立変数が有意な効果を示したとしても、それは純粋に第1水準の効果だけを示しているのではない。第1水準と第2水準の混合効果を示

⁴ **partner effect** と **mutual influence** の違いはわかりにくいかもしれないが、注目している同一変数内での直接的な影響が **mutual influence** であり、別の原因変数を介した影響が **partner effect** である。

していることになる（少なくともその可能性があることになる）。

3.2 NFRJ のマルチレベル分析で留意が不要な点、必要な点

では、以上のような枠組みに沿って、ダイアド集積型家族調査としての NFRJ にマルチレベル・モデリングを適用する際の方法を考察しよう。まず、マルチレベル分析の前提となる非独立性の扱いについて考える。先に記したように、ダイアドデータにおける非独立性には正の相関によるものだけでなく負の相関によるものも含まれていることに注意が必要である。たとえば、NFRJ で同じ回答者の複数の子ダイアドについて関係満足度を分析する場合、子ダイアド間には満足度の正の相関も負の相関も考えられる。ここで問題になってくるのは、ICC の扱い方である。ICC の根本的な意味は、グループ内での相関であるが、実際的には全分散のうちグループの違いで説明できる分散の割合として算出・利用されることが多い。また、ある独立変数群を投入した上での ICC は、**partial ICC** と呼ばれ、投入済みの独立変数によって説明できる分散を除いた ICC として評価される。ただし、このような ICC の算出・解釈ができるのは、ICC が正の相関であることを前提にしている。NFRJ の分析の場合には、負の相関も一定程度ありえることを考慮しなければならない。とはいえ、ほとんどの場合は、同じ家族・親族に属することによる類似性の方が勝り、全体としては正の相関が強く出ることであろう。したがって、目安として ICC で共通性の程度を読み取ることに問題は無いが、モデル間で ICC を比較してその変動から説明力を読み取るような利用は危険と考えられる。

また、NFRJ ではカテゴリカル変数が多く、従属変数をダミー変数とする二項ロジットモデル（ロジスティック回帰）等によるマルチレベル・モデリングが多くなると予想される。従属変数がダミー変数の場合には、ICC を算出することはできない。ただし、数理上は、ダミー変数の背後に潜在的な連続変数があり、それが閾値を超えるかどうかでダミー変数の 0/1 が決まっていると想定するので、その潜在変数を従属変数に見立てて、ICC を算出することがある。Stata のように一部の統計ソフトは、この数値を出力するが（Rabe-Hesketh and Skrondal 2008）、やはり擬似的な ICC なので注意が必要である。全体的に言って、NFRJ の分析で ICC の変動を細かく検討することは危険を伴う。

次に、大枠となる **null model**、**random intercepts model**、**random slopes models** の選択について考える。結論としては、NFRJ では特殊な場合を除いて **random intercepts model** のみを考えるべきであろう。非独立性がある限り、**null model** が妥当でなく、少なくとも **random intercepts** を想定しなければならないことは自明である。一方、**random slopes models** の検討を除外するのは、1 人の回答者に属するダイアド数が少ないことによる。通常の場合に多くの独立変数の係数（傾き）を推定するためには、それにみあったケース数が必要となることと同様に、マルチレベル分析においても、グループごとにそれぞれ傾きの異なる回帰式を推定するためには、それぞれのグループに一定以上のケース数が必要になる。p 個の独立変数の傾きをすべてランダムにするためには、ケース数 k が少なくとも p+2 以上なけ

ればならない (Kenny et al. 2006: 279)。NFRJ の場合、たとえば子ダイアドやきょうだいダイアドに注目する場合、1 グループのケース数は最大 3 であり、平均的には 2 ケースを切る場合もある。たった 1 つの傾きすらランダムにするには不十分である。random slopes models が有効な状況は、種類の異なるダイアドをまとめて分析するような特殊な分析でケース数がある程度確保される場合に限られるだろう (さまざまな種類のダイアドの中で関係満足度の決まり方がどう異なるのかを分析するような場合)。そうでない場合の分析が完全に不可能ということではないが、十分な検出力で安定した結果が得られない危険が高い。NFRJ の分析は、基本的に random intercepts model で考えるべきである。

マルチレベル分析のテキストでは、しばしばセンタリングの必要性やその方法に多くの解説をあてている。センタリングとは、独立変数の値から全体平均 (grand mean) やグループ内での平均 (group mean) を引いて、平均が 0 になるように変換することである。マルチレベル分析では、このセンタリングの仕方によって、分析結果が読み取りやすくなるだけでなく、センタリングしない場合と比べてまったく意味の異なるモデルを推定していることになったりする。しかし、センタリングが重要な意味を持つのは、基本的に random slopes models を扱う場合に限られる。NFRJ の分析では、上で述べたとおり、ほとんどの場合において random slopes models は扱うことができないので、センタリングの問題に神経をとがらせる必要はない。わかりやすさの意味でも、そのままの値で分析を進めればよい。

逆に、NFRJ のマルチレベル分析でとくに注意すべき点は、第 1 水準での独立変数の投入の仕方である。先に示したように、ダイアドデータ分析においては、第 1 水準の独立変数の多くが mixed variables なので、純粋に第 1 水準の効果を示すのではなく、第 2 水準の効果の一部が混ざっている可能性を考慮すべきである。NFRJ のようにグループごとのケース数が少ない場合には、とくにこの意味が大きい。たとえば、それぞれの子ダイアドの関係満足度を従属変数とするマルチレベル分析で、子どもとの居住距離を独立変数とする場合を考えよう。距離が短いことが満足度にプラスの効果があるという分析結果が出たとしても、この効果は mixed であることに注意しなければならない。つまり、まさにその子どもが近くに住んでいるという純粋に第 1 水準の効果と、(他の子どもも含めて) 回答者が全体として子どもと近く住んでいるという第 2 水準の効果が混ざっていると考えられる。これを弁別するためには、「すべての子どもとの平均距離」のように第 2 水準の効果を表す変数を作成し、これを同時に投入してやればよい。平均距離でよいのか、たとえばいちばん近い子どもとの距離のように効果が最大と考えられるもので代表させるべきかは、理論に依存する。

その一方で、子ども同士の間にも交流があるので、他のダイアドによる効果にも注意しなければならない。つまり、(当該の子ダイアドを除いた) 他の子どもとの距離が効果を持つ可能性がある。たとえば、近所に住んでいる子ども (あるいは同居の子ども) がいると、他の子どもとの関係が不満になる、というような理論が成り立つ場合である。この場合も、他の子どもとの平均距離 (当該の子どもを除いているという点で、前の段落の平均距離と

は異なる) や、他の子どもの中でのもっとも近い距離といった変数を作成し、分析に投入すればその効果を検討することができる。

ただし、NFRJ では1つのグループに属するダイアドの数が少ないので、上に示したような、第1水準のそのままの変数、全体の平均、他のダイアドの平均の3者間には強力な関連性があり、同時に投入すると多重共線性により安定的な結果は得られないであろう。統計的な結果と理論的な解釈可能性の両面から最適なモデルを模索しなければならない。NFRJ データのマルチレベル分析においてもっとも困難で、また意味のある過程は、この独立変数の構成にあると考えられる。

4. 分析事例

4.1 分析対象と基本モデル

以上のような考察に基づいたマルチレベル分析の具体的な事例を示そう。この節では、NFRJ08における世代間援助行動をダイアド集積型のマルチレベル・モデリングで分析する。NFRJ08 (National Family Research of Japan 2008; 第3回全国家族調査) は2009年1~2月に実施された訪問留置調査で、2008年12月31日現在で28~72歳の全国の男女9,400人を対象にしているが(回収数5,203、回収率55.35%)、ここでは分析対象を明確にするために、回答者を中期親子の親と想定して分析サンプルを限定する。具体的な分析サンプルは、回答者(親)が65歳以下で、かつ子が25歳以上の親子ダイアドで、使用する変数が欠損する不完全ケースを除いた結果、1,268の回答者による2,397の親子ダイアドが分析対象となった。1人の回答者(親)が平均1.89のダイアドを持つ。

NFRJ08では、それぞれの子どもの間で、過去1年間に、経済的援助、実践的援助、情緒的援助の授受があったかどうかを尋ねており、各ダイアドについて $3 \times 2 = 6$ 通りの援助の有無が調べられている⁵。ここでは、これら6つの援助行動のダミー変数を従属変数として、二項ロジット(ロジスティック回帰)のマルチレベル・モデリングで分析を行う。先の考察に基づき、ランダム切片モデル(random intercept models)のみを検討し、傾きのランダム性は認めない。つまり、表面上は、切片のランダム誤差が算出される以外には、通常のロジスティック回帰分析と同じように分析結果が得られる。分析には、ランダム切片モデルの扱いが容易なStata 11のxtlogitコマンドを用いた⁶。データはあらかじめ、回答者単位か

⁵ 経済的援助は、金銭的な援助(小遣い、仕送り、贈与など)の授受で、30万円以上と30万円未満に分けて選択肢が設けられているが、ここではこれらは区別しない。実践的援助は、看病や家事・育児の授受であり、情緒的援助は相談相手になること、相談にのってもらうことである。NFRJ08では、若年調査票(28~47歳対象)のみ、子ダイアドとの援助関係が省かれている。そのため、ここでの分析でも47歳以下の親と25歳以上の子どもからなる親子ダイアドは分析対象外となっている。

⁶ 他に、HLM 6、Mplus 6、SPSS 18での分析を検討したが、少なくともランダム切片モデルに限れば、STATA 11がデータハンドリングや記述統計の出力等を含めて総合的に優れていた。Stataにおけるマルチレベル分析の方法については、Rabe-Hesketh and Skrondal (2008)を参照した。

らダイアド単位に変換している⁷。

援助の有無を説明する独立変数は、子側と親側のそれぞれから与えられる。ここでは、親が回答者なので、子側が第1水準、親側が第2水準のマルチレベル・モデリングである。子側の要因としては、性別（男）、年齢、同居、5歳以下の孫の有無、6歳以上の孫の有無、大卒か否か、を投入した⁸。年齢以外はダミー変数である。親側の要因としては、性別（男）、年齢、健康状態（1～5点で5が「たいへん良好」）、子ども数、有配偶か否か、大卒か否かを投入した。年齢、健康状態、子ども数以外はダミー変数である。先の考察に基づいて、いずれの変数も平均値によるセンタリングは行っていない。各変数の記述統計は、表1のとおりである。

表1 使用変数の記述統計

従属変数		独立変数 (Level-1 : 子)		独立変数 (Level-2 : 親)	
子どもへの援助 (経済)	27.8%	性別 (男)	52.3%	性別 (男)	40.0%
子どもへの援助 (実践)	28.8%	年齢	31.06 (4.07)	年齢	58.59 (3.68)
子どもへの援助 (情緒)	55.0%	同居	33.7%	健康状態	3.60 (0.87)
子どもからの援助 (経済)	13.3%	5歳以下の孫	29.4%	子ども数	2.44 (0.71)
子どもからの援助 (実践)	26.0%	6歳以上の孫	17.2%	有配偶	88.4%
子どもからの援助 (情緒)	31.4%	大卒	34.3%	大卒	12.5%

注：ダミー変数の記述統計は比率、連続変数は平均（括弧内は標準偏差）

独立変数による説明モデルに先だって、6つの援助行動のそれぞれについて、ICCを算出したところ、図2のようになった。先の考察で示したとおり、負の相関を考慮しなければならないNFRJのダイアドデータでは、この数値を「親が共通であることで説明できる従属変数の分散の割合」と解釈することは危険であるが、少なくともいずれの援助もICCが

⁷ 手続き的には、配偶者やきょうだいを含めた11のダイアド全てについて、下にあげる変数を第1水準の変数として作成し、1人の回答者ケースを11のダイアドケースに変換した上で、今回の分析に必要な子ダイアドのケースおよび使用変数を抜き出した。変換されたダイアド単位のデータは、親子ダイアド以外の分析にも汎用的に利用可能である。

作成した下位水準の変数：ddrlt「ダイアドの種類」（配偶者、実親、義親、子、きょうだい）、ddodr「順位（子・きょうだいのみ）」、ddsex「性別」、ddalive「生死」、ddage「年齢」、ddaged「回答者との年齢差」、ddmrg「婚姻状態」、ddjob「仕事の有無」、ddrsd「居住関係」、ddedu「学歴」、ddccno「子の有無（なし）」、ddcc00「子の有無（0～5歳）」、ddcc06「子の有無（6～17歳）」、ddcc18「子の有無（18歳以上）」、ddent「会話頻度」、ddrvfn「援助を受ける（経済）」、ddrvem「援助を受ける（情緒）」、ddrvpr「援助を受ける（実践）」、ddgvfn「援助を与える（経済）」、ddgvem「援助を与える（情緒）」、ddgvpr「援助を与える（実践）」、ddevl「関係の評価（配偶者は満足度で代替）」、ddfam「家族の一員か」、ddfams「その配偶者が家族の一員か」

⁸ 今回の分析では、世代間援助関係にとって重要な変数である子側の出生順位（長男か否かなど）を投入していない。これは、後で「同居」変数の投入の仕方について論じるのと同様に、出生順位の変数をどのように投入すべきかは、慎重に検討される必要があるためである。ここでは、分析事例としてのわかりやすさを優先してモデルを構成しているため、後に示す分析結果を世代間援助関係のクリティカルな知見として扱うべきではないことをお断りしておく。

十分に大きく、マルチレベル・モデリングでこれらの変数を扱うことが妥当であることが読み取れる。また、相対的に子どもへの実践的援助において ICC が小さく、親が子どもの家事を手伝うことは、他の援助に比して、子どもごとに違いが出やすい（つまり、子側の要因で援助の有無が左右される）ことが読み取れる。

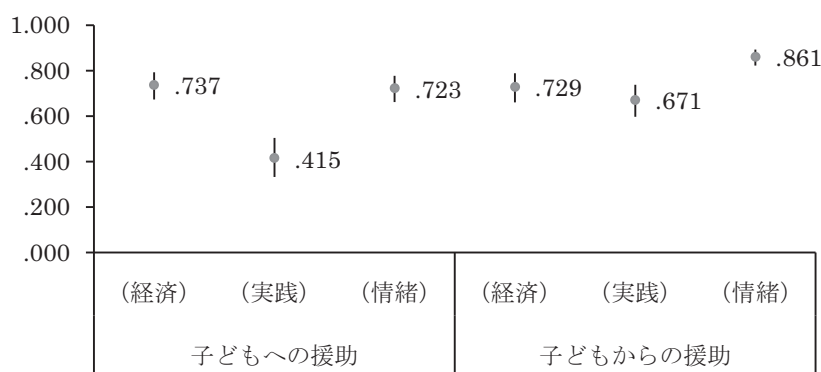


図2 各援助行動の ICC (95%信頼区間)

4.2 マルチレベル・モデリング以外との対比

6つの援助行動の有無について、それぞれランダム切片のマルチレベル・モデリングで分析した結果が、表 2-a～f である。マルチレベルを設定したことの意義をわかりやすく示すために、別途 2つの参考モデルを並置している。最初の列がマルチレベル・モデリングの結果であり、2番目の列は第1子との関係にケースを限定して分析した結果である（1回答者につき1人の子どもなので、マルチレベルではない）。3番目の列は、全てのダイアドケースについて、マルチレベル・モデリングではなく通常の二項ロジットモデルで分析した結果（つまり、ダイアドケース間に独立性が成り立っているとみなした場合の結果⁹）である。1番目と3番目のモデルでケース数が2,397なのに対して、2番目のモデルではケース数が1,260となる。

各独立変数の有意性だけに注目すると、どの援助行動についても、マルチレベル・モデリングの結果は、2つの参考モデルと比べて大きな違いを示すものではない。しかし、それぞれの有意な独立変数について係数のサイズに注目すると、マルチレベルの結果は多くの係数でそのサイズが大きくなっている。つまり、援助行動の要因と目される変数の影響力の大きさがより鮮明に示されていることがわかる。たとえば、表 2-a（子どもへの経済的援助）では、未就学児（5歳以下の孫）の存在が援助の発生率を有意に高めることが3つのモデルで共通に示されているが、2つの参考モデルではそのオッズ比は1.5倍程度であるのに対して、マルチレベルでは2.24倍である。同じように、親が大卒であることの効果は、

⁹ いうまでもなく、独立性の仮定は明らかに誤りであり、この分析モデルは統計的にも理論的にも不適切である。あくまで、マルチレベル・モデリングとの間で回帰係数の対比を行うための参考モデルである。

参考モデルでは 2 倍前後のオッズ比であるのに対して、4.15 倍もの数値を示している。
この傾向は、援助の種類や方向（授受）が異なっても同じである。

表 2-a 子どもへの援助（経済）

	マルチレベル・モデリング				参考1：第1子との関係のみ分析				参考2：すべて独立ケースとみなした分析			
	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値
Level 1 子												
性別（男）	-.03	.17	.97	.841	.02	.13	1.02	.856	.01	.10	1.01	.916
年齢	-.11	.03	.90	.001 **	-.08	.03	.92	.002 **	-.07	.02	.94	.000 ***
同居	-.31	.20	.73	.109	-.13	.15	.87	.375	-.19	.11	.82	.074 +
5歳以下の孫	.81	.20	2.24	.000 ***	.46	.15	1.58	.002 **	.43	.11	1.54	.000 ***
6歳以上の孫	.78	.26	2.19	.003 **	.74	.19	2.10	.000 ***	.45	.14	1.56	.001 **
大卒	-.25	.20	.78	.224	-.07	.15	.93	.624	-.08	.11	.93	.470
Level 2 親												
性別（男）	-.20	.26	.82	.438	-.23	.15	.79	.116	-.11	.11	.90	.316
年齢	-.09	.04	.91	.014 *	-.03	.02	.97	.279	-.04	.02	.97	.034 *
健康状態	.04	.13	1.04	.792	.03	.07	1.03	.731	.03	.05	1.03	.564
子ども数	-.22	.16	.80	.171	-.13	.09	.88	.150	-.07	.07	.94	.323
有配偶	.02	.37	1.02	.951	-.12	.20	.89	.561	.01	.15	1.01	.972
大卒	1.42	.36	4.15	.000 ***	.72	.19	2.05	.000 ***	.64	.14	1.90	.000 ***
切片	6.80	2.02			3.24	1.13			3.05	.81		
ランダム誤差 SD	2.98											

表 2-b 子どもへの援助（実践）

	マルチレベル・モデリング				参考1：第1子との関係のみ分析				参考2：すべて独立ケースとみなした分析			
	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値
Level 1 子												
性別（男）	-1.33	.18	.27	.000 ***	-.75	.15	.47	.000 ***	-.86	.11	.43	.000 ***
年齢	-.01	.03	.99	.740	-.02	.03	.98	.471	-.01	.02	.99	.621
同居	1.44	.20	4.20	.000 ***	.77	.17	2.17	.000 ***	.88	.12	2.42	.000 ***
5歳以下の孫	3.53	.27	34.00	.000 ***	2.06	.17	7.85	.000 ***	2.09	.12	8.10	.000 ***
6歳以上の孫	1.53	.24	4.61	.000 ***	1.12	.20	3.05	.000 ***	.91	.15	2.48	.000 ***
大卒	.19	.19	1.21	.320	.24	.16	1.27	.141	.19	.12	1.21	.100
Level 2 親												
性別（男）	-1.58	.24	.21	.000 ***	-1.11	.17	.33	.000 ***	-.90	.12	.41	.000 ***
年齢	-.03	.03	.97	.351	-.01	.03	.99	.621	-.02	.02	.98	.387
健康状態	.25	.11	1.28	.029 *	.10	.08	1.11	.213	.13	.06	1.14	.024 *
子ども数	-.23	.14	.80	.092 +	-.17	.10	.84	.086 +	-.14	.07	.87	.064 +
有配偶	-.14	.30	.87	.632	.03	.22	1.03	.893	-.11	.16	.90	.494
大卒	.56	.32	1.75	.076 +	.38	.23	1.46	.093 +	.27	.17	1.32	.111
切片	-.65	1.65			-.05	1.24			-.33	.90		
ランダム誤差 SD	2.19											

表 2-c 子どもへの援助（情緒）

	マルチレベル・モデリング				参考1：第1子との関係のみ分析				参考2：すべて独立ケースとみなした分析			
	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値
Level 1 子												
性別（男）	-1.05	.17	.35	.000 ***	-.53	.12	.59	.000 ***	-.47	.09	.62	.000 ***
年齢	-.12	.03	.89	.000 ***	-.06	.02	.94	.004 **	-.07	.02	.93	.000 ***
同居	.29	.18	1.34	.107	.41	.14	1.50	.003 **	.22	.10	1.24	.026 *
5歳以下の孫	.56	.18	1.75	.003 **	.24	.14	1.27	.087 +	.25	.10	1.29	.010 *
6歳以上の孫	.07	.24	1.07	.770	.13	.17	1.13	.462	.10	.13	1.10	.444
大卒	-.01	.19	.99	.947	.18	.13	1.20	.174	.12	.10	1.12	.226
Level 2 親												
性別（男）	-1.10	.25	.33	.000 ***	-.49	.13	.61	.000 ***	-.49	.10	.61	.000 ***
年齢	-.06	.04	.94	.090 +	-.02	.02	.98	.296	-.02	.02	.98	.231
健康状態	-.03	.13	.97	.812	-.04	.07	.96	.567	-.03	.05	.97	.550
子ども数	-.23	.16	.80	.150	-.11	.08	.90	.195	-.10	.06	.90	.089 +
有配偶	-.68	.37	.51	.064 +	-.40	.19	.67	.038 *	-.31	.14	.73	.022 *
大卒	1.50	.37	4.49	.000 ***	.60	.19	1.82	.002 **	.69	.14	1.98	.000 ***
切片	9.63	2.03			4.53	1.07			4.27	.76		
ランダム誤差 SD	3.02											

表 2-d 子どもからの援助 (経済)

	マルチレベル・モデリング				参考1: 第1子との関係のみ分析				参考2: すべて独立ケースとみなした分析			
	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値
Level 1 子												
性別 (男)	-.46	.23	.63	.051 +	-.18	.17	.84	.299	-.29	.13	.75	.026 *
年齢	.07	.04	1.07	.093 +	.05	.03	1.05	.130	.02	.02	1.02	.268
同居	2.26	.28	9.56	.000 ***	1.24	.19	3.45	.000 ***	1.29	.14	3.62	.000 ***
5歳以下の孫	-.71	.30	.49	.017 *	-.16	.22	.86	.485	-.45	.17	.64	.010 *
6歳以上の孫	-.31	.35	.74	.382	-.33	.26	.72	.192	-.29	.21	.75	.162
大卒	.27	.26	1.31	.313	.06	.19	1.06	.739	.11	.14	1.12	.426
Level 2 親												
性別 (男)	-.93	.34	.40	.007 **	-.38	.20	.69	.054 +	-.53	.15	.59	.000 ***
年齢	.05	.05	1.05	.320	.03	.03	1.03	.309	.04	.02	1.04	.105
健康状態	-.29	.17	.74	.081 +	-.15	.09	.86	.113	-.17	.07	.84	.015 *
子ども数	.28	.20	1.32	.163	.15	.11	1.17	.169	.20	.09	1.22	.019 *
有配偶	.17	.47	1.19	.714	.15	.27	1.16	.576	.07	.20	1.07	.728
大卒	-.13	.49	.88	.793	-.14	.28	.87	.632	-.04	.22	.97	.869
切片	-9.05	2.67			-5.40	1.53			-4.83	1.13		
ランダム誤差 SD	3.20											

表 2-e 子どもからの援助 (実践)

	マルチレベル・モデリング				参考1: 第1子との関係のみ分析				参考2: すべて独立ケースとみなした分析			
	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値
Level 1 子												
性別 (男)	-2.63	.28	.07	.000 ***	-1.30	.16	.27	.000 ***	-1.21	.11	.30	.000 ***
年齢	-1.10	.04	.91	.014 *	-.05	.03	.95	.050 +	-.03	.02	.97	.091 +
同居	3.92	.36	50.52	.000 ***	1.96	.17	7.13	.000 ***	1.82	.12	6.17	.000 ***
5歳以下の孫	.36	.25	1.44	.141	.20	.18	1.22	.284	.21	.13	1.24	.093 +
6歳以上の孫	.36	.32	1.43	.267	-.05	.23	.95	.827	-.05	.16	.95	.738
大卒	.26	.25	1.30	.298	.14	.16	1.15	.391	.12	.12	1.13	.312
Level 2 親												
性別 (男)	-1.14	.32	.32	.000 ***	-.62	.17	.54	.000 ***	-.47	.12	.63	.000 ***
年齢	-.05	.04	.96	.312	.00	.03	1.00	.874	-.03	.02	.98	.169
健康状態	-.58	.16	.56	.000 ***	-.18	.08	.84	.031 *	-.29	.06	.75	.000 ***
子ども数	.19	.19	1.20	.333	.01	.10	1.01	.951	.10	.07	1.11	.161
有配偶	-.70	.43	.49	.105	-.36	.22	.69	.104	-.32	.16	.73	.045 *
大卒	.38	.44	1.46	.384	.16	.23	1.17	.498	.19	.17	1.20	.275
切片	5.15	2.40			1.24	1.27			2.35	.90		
ランダム誤差 SD	3.45											

表 2-f 子どもからの援助 (情緒)

	マルチレベル・モデリング				参考1: 第1子との関係のみ分析				参考2: すべて独立ケースとみなした分析			
	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値	係数	SE	オッズ比	p値
Level 1 子												
性別 (男)	-1.43	.23	.24	.000 ***	-.51	.13	.60	.000 ***	-.46	.10	.63	.000 ***
年齢	.04	.04	1.04	.314	.01	.02	1.01	.609	.00	.02	1.00	.958
同居	1.35	.24	3.87	.000 ***	.65	.15	1.92	.000 ***	.64	.11	1.90	.000 ***
5歳以下の孫	.63	.24	1.87	.009 **	.40	.15	1.49	.008 **	.36	.11	1.43	.001 **
6歳以上の孫	-.12	.30	.89	.703	-.04	.19	.96	.830	-.04	.14	.97	.802
大卒	.67	.25	1.96	.007 **	.32	.14	1.37	.027 *	.37	.10	1.45	.000 ***
Level 2 親												
性別 (男)	-2.97	.42	.05	.000 ***	-.98	.15	.38	.000 ***	-1.11	.11	.33	.000 ***
年齢	-.03	.05	.97	.562	.00	.03	1.00	.983	.00	.02	1.00	.855
健康状態	-.58	.19	.56	.002 **	-.25	.07	.78	.001 **	-.26	.05	.77	.000 ***
子ども数	.02	.22	1.02	.913	.01	.09	1.01	.903	.01	.07	1.01	.883
有配偶	-1.35	.50	.26	.007 **	-.35	.19	.70	.067 +	-.48	.14	.62	.000 ***
大卒	.28	.52	1.32	.592	.02	.21	1.02	.908	-.02	.16	.98	.911
切片	2.31	2.72			.19	1.12			.90	.81		
ランダム誤差 SD	4.28											

注: *** p<.001, ** p<.01, * p<.05, + p<.10

マルチレベル・モデリングのこのような推定結果は、同じ親に属する複数の親子ダイアドを直接的かつ適切に比較したことの成果である。つまり、5歳以下の孫の効果は、同じ親について、5歳以下の孫がいる子どもといない子どもを直接比較することで知ることができ、また、親が大卒であることの効果は、さまざまな面で異なる子どものいずれもが同じ大卒の親を持つと援助を受けやすいことを観察することで知ることができる。第1子のダイアドのみを分析した参考モデル1では、統計的には適切なモデルであっても、肝心の比較データが取り除かれており、理論的に不適切なモデルということができる。すべてのダイアドを独立ケースとみなした参考モデル2では、同じ回答者（親）に属する親子ダイアドの類似性（非独立性）が無視されており、統計的に不適切なモデルである。理由は異なれども、それぞれの不適切な2つの参考モデルが似たような係数を推定していることは興味深い。

4.3 子どもへの経済的援助における mixed variables の検討

先の考察で、ダイアド集積型の NFRJ データでは、第1水準の変数のほとんどが mixed variables であるという認識が重要であることを示した。また、そのために、実践面では、第1水準の変数を加工した変数を慎重に投入し、mixed variables の効果を第1水準と第2水準に分割することを試みる必要があることを示した。ここの分析例では、子側の独立変数はすべて mixed variables であり、それぞれの変数が有意であったとしても、「子が〇〇である」ことの効果と「親が〇〇な子を持っている」ことの効果のいずれなのかを単純には判断できない。また、同時に「他に〇〇な子がいる」ことによるダイアド間効果の可能性も考えなければならない。そこで、この節では、子どもへの経済的援助の分析を例に、第1水準の変数を加工した変数の投入で、それぞれの効果の判別を試みる。

表3は、子どもへの経済的援助を従属変数とした、4つのマルチレベル・モデリングの分析結果である。モデル1は、前節で示したモデルと同じである。第1水準の独立変数のうち、量的変数である年齢は混合効果の検討が難しいので、その他のダミー変数の効果に注目することにする¹⁰。5歳以下、6歳以上の孫の存在がそれぞれ援助の発生率を有意に高めており、また、有意水準には達していないものの、同居が援助の発生率を低める傾向が見られる。

¹⁰ 後で示すように、ダミー変数の場合は、「他子との同居の有無」のようにダイアド間効果を示す変数を作成することができる。量的変数の場合には、たとえば、「他子の年齢の平均値」「他子の年齢の最高値」といった変数がこれに相当すると考えられるが、一人っ子の場合には、これらの変数を作成することができず、同じ方法で検討のステップを進めることはできない。この問題への対応として、変数を全体平均値（grand mean）でセンタリングし、「他子の年齢の平均値」を作成する場合は、他子の年齢が全体平均と一致する場合も、他子がそもそもいない場合も同じように0を投入するという考え方が成り立つ。ただし、この問題はまだ十分に検討していないので、本稿では量的変数の場合の扱いを保留し、ダミー変数のみに注目している。

表3 子どもへの援助（経済）のモデル検討

	モデル1 (再掲)			モデル2			モデル3			モデル4		
	係数	SE	p値	係数	SE	p値	係数	SE	p値	係数	SE	p値
Level 1 子												
性別 (男)	-.03	.17	.841	-.05	.17	.772	-.04	.17	.801	-.06	.17	.736
年齢	-.11	.03	.001 **	-.11	.03	.001 **	-.11	.03	.001 **	-.11	.03	.001 **
同居	-.31	.20	.109	-.44	.21	.037 *	-.17	.22	.418			
5歳以下の孫	.81	.20	.000 ***	.76	.22	.000 ***	.80	.20	.000 ***	.84	.19	.000 ***
6歳以上の孫	.78	.26	.003 **	.76	.28	.007 **	.79	.26	.003 **	.81	.26	.002 **
大卒	-.25	.20	.224	-.24	.20	.245	-.24	.20	.242	-.22	.20	.279
他子と同居				-.38	.22	.078 +						
他子に5歳以下の孫				-.07	.22	.740						
他子に6歳以上の孫				-.07	.27	.794						
Level 2 親												
性別 (男)	-.20	.26	.438	-.20	.26	.450	-.20	.26	.428	-.21	.26	.424
年齢	-.09	.04	.014 *	-.10	.04	.009 **	-.10	.04	.009 **	-.10	.04	.008 **
健康状態	.04	.13	.792	.04	.13	.751	.04	.13	.742	.04	.13	.750
子ども数	-.22	.16	.171	-.14	.17	.435	-.17	.17	.300	-.15	.16	.346
有配偶	.02	.37	.951	.02	.38	.967	.01	.37	.985	.00	.37	.992
大卒	1.42	.36	.000 ***	1.41	.37	.000 ***	1.41	.36	.000 ***	1.41	.36	.000 ***
いずれかの子と同居							-.44	.27	.110	-.53	.25	.035 *
切片	6.80	2.02		7.33	2.10		7.34	2.06		7.29	2.05	
ランダム誤差 SD	2.98			2.98			2.98			2.97		
LL	-1223.1			-1221.6			-1221.8			-1222.2		

注：*** p<.001, ** p<.01, * p<.05, + p<.10

そこで、これらの独立変数について、親が他子との間に同様の関係を持っているかどうかというダミー変数を作成し、ダイアド間効果の有無を検討したのがモデル2である。結果、他子ダイアドに孫が存在することは、まったく効果を持たないことがわかった。他子の効果の完全な否定は、その集合でもある親水準の効果（いずれかの子ダイアドに孫がいることが全ての子ダイアドに同様の効果を持つこと）の否定でもある。孫の効果に関しては、単純に第1水準の効果（孫の存在する子どもには、経済的援助が発生しやすい）とみなしてよさそうである。一方で、「他子との同居」については10%水準で有意な効果が見られた。つまり、当該ダイアドの子と同居していなくとも、他の子と同居している親からは同様に経済的援助が得にくいことを示している。

モデル2では「他子との同居」と「(当該ダイアドの) 子との同居」がほぼ同様の効果を示しているため、これは第2水準（親水準）の効果を示している可能性が高い。そこで、「いずれかの子との同居」を示す親水準の変数を明確に投入したのがモデル3である。この変数は統計的には有意になっていないが、モデル3で「(当該ダイアドの) 子との同居」の効果が大きく減退しているため、やはり、とくに当該ダイアドの子どもと同居していることが重要なのではなく、親がいずれかの子と同居していることの方が重要であることがわかる。

最終的に、同居に関して「いずれかの子との同居」だけを残したモデル4では、この効果が統計的にも有意であることが確認できる。この結果の正確な意味を解釈するためにはさらに追加の分析が必要だが、恐らくは、65歳以下という比較的若い段階で成人子（25歳以上の子ども）と同居する必要がある親は、あまり経済的に余裕がなく、子どもに経済的援助を与えることが少ないのではないかと予想される。

このように、1人の回答者に属するダイアドが少数である NFRJ データでは、mixed variables である第1水準の独立変数の効果を解釈する際に慎重な検討が必要になる。積極的な言い方をすれば、1人の回答者につき1つの親子ダイアドしか観察しない通常のデータでは不可能であった領域に、NFRJ データは分析を進めることを可能にする。従来のデータでは、子側の効果と親側の効果を統計的に区別することができなかった（そして、多くの場合は不十分な検討で子側の変数は子側の効果と単純に捉えられてきた）が、NFRJ データのマルチレベル・モデリングでは、その検討が可能になる。

5. まとめ

本稿では、ダイアド集積型家族調査である NFRJ をマルチレベル分析で有効に活用する方法を考察してきた。マルチレベル分析の応用範囲は広く、その方法をすべて適切に習得し利用することは必ずしも容易ではないが、こと NFRJ データの分析に限れば、ほとんどの場合ランダム切片モデルだけに意識を集中すればよく、そのため変数のセンタリングにも気を遣う必要がないことを示した。逆に、多くのマルチレベル分析ではあまり重視されないが、NFRJ データの分析では重要になるポイントがあることも示した。ICC に正の相関だけでなく負の相関もあることを考慮しなければならず、そのため ICC を第2水準のグループで説明できる分散の割合と捉えることは危険である。また、第1水準（ダイアド水準）の独立変数がほとんど mixed variables であり、本当に個別のダイアドの水準の効果なのか、それとも混入した第2水準の効果なのか、はたまた他のダイアドからのダイアド間効果が混じっているのかといった検討が必要になることも示した。

事例として、NFRJ08 データから中期親子の援助行動の分析を行い、マルチレベル・モデリングの結果が、従来よりも非常に鮮明に援助行動の規定要因を表せることがわかった。また、子どもへの経済的援助について、子側の mixed variables の検討を行い、子どもとの同居の効果が、当該ダイアドの子との同居に限ったものではなく、いずれかの子どもと同居しているという親側の要因であることを示した。

ダイアド集積型家族調査のマルチレベル・モデリングは、比較すべき複数のダイアドを直接的に比較する、という極めて当たり前の分析志向を具現化している。方法論的な検討と統計分析ソフトの対応が進んだ現在、NFRJ データにこの手法を適用することは、多くの有意義な研究成果をもたらすものと期待できる。

[文献]

Kenny, David A., Deborah A. Kashy, and William L. Cook, 2006, *Dyadic Data Analysis*, New York: Guilford Press.

Kenny, David A. and Deborah A. Kashy, 2011, "Dyadic Data Analysis Using Multilevel Modeling," Joop J. Hox, and J. Kyle Roberts eds. *Handbook of Advanced Multilevel Analysis*, New York: Routledge, 335-370.

- Lee, Kristen Schultz, 2009, "Competition for Resources: A Reexamination of Sibship Composition Models of Parental Investment," *Journal of Marriage and Family*, 71: 263-277.
- Luke, Douglas A., 2004, *Multilevel Modeling*, Thousand Oaks: Sage Publications.
- Rabe-Hesketh, Sophia and Anders Skrondal, 2008, *Multilevel and Longitudinal Modeling Using Stata*, Second Edition. College Station: Stata Press.
- Snijders, Tom and Roel Bosker, 1999, *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London: Sage Publications.
- Teachman, Jay and Kyle Crowder, 2002. "Multilevel Models in Family Research: Some Conceptual and Methodological Issues," *Journal of Marriage and Family*, 64: 280-294.
- 渡辺秀樹, 2001, 「NFR98 の思想」 嶋崎尚子編『家族生活についての全国調査 (NRJ98) 報告書 No.2-7 家族と職業』日本家族社会学会全国家族調査 (NFR) 研究会, 79-88.

Methods to Apply Multilevel Modeling to NFRJ Survey Data: Effective Use of Multiple Dyadic Family Surveys

Tokio YASUDA

Osaka University of Commerce

The purpose of this paper is to consider the concrete methodological matters for applying multilevel modeling to NFRJ survey data. NFRJ is a multiple dyadic family survey, namely it is composed of eleven dyads related to individual respondents. It is effective to apply multilevel modeling in which level-1 units are dyads and level-2 units are respondents.

I examined methods for the modeling of NFRJ, based on methodological studies on dyadic data analysis. The results are the followings. First, you should test random intercept models and not random slope models because the group size of dyads that belong to single respondents is small in NFRJ. Second, negative correlation between dyads is possible on NFRJ, therefore it is necessary to interpret ICC carefully. Third, it is difficult to interpret an effect of a level-1 predictor because level-2 effect is possible to be mixed in it intensively. Finally, an analysis of the intergenerational support illustrated that multilevel modeling of NFRJ should show a clear and meaningful results.

Key words and phrases: multilevel analysis, dyadic data, family survey, methodology

