

Donnellan,E., Usami,S., Murayama,K., (in press). Random item slope regression: Examining both similarities and differences in the association with individual items. *Psychological Methods*.

報告者：宇佐美慧

心理尺度やテストを用いた心理学研究において、複数の項目に基づく尺度得点（項目和得点）を従属変数として扱い、従属変数と（対象の年齢などの、単一の項目である）独立変数との関係を推測するための回帰分析が実行されることは多い。他にも、尺度得点ではなく、項目群の背後に潜在変数（共通因子）を仮定して、潜在変数と（単一の項目である）独立変数との関係を構造方程式モデリング（SEM）により推測するアプローチがとられることも多い。しかし、これらの心理学研究で頻繁になされる分析実践は、「独立変数と各尺度項目との関連が等質である」、または「独立変数と各尺度項目との関連の強さが、因子負荷に比例した形で構造化されている」といった強い仮定の下で成立している。つまり、独立変数が各項目に異なる影響を及ぼしうる可能性、言い換えればランダムな項目傾き（random item slopes）が存在する可能性は考慮されていない。

本研究ではこのような分析実践の対処として、ランダムな項目傾きを含む線形混合効果モデル（random item slope regression）を提示し、特に変数間の関係の強さを反映する回帰係数の点推定値・標準誤差および解釈の観点から、既存のアプローチとの異同について、概念的・数理的な観点から説明した。このモデルはいわば、利用されている当該の項目が項目ユニバース（item universe）からのサンプルで、項目に応じて係数がランダムに変動することを変量効果として扱うモデルである。その結果として、モデル内の（例えば項目を隔てた係数の平均などの）推定値も項目ユニバースに一般化可能な量として解釈され、特に研究で扱う構成概念の幅が広く異質性の高い項目群を対象とする場合、解釈上より妥当な結論を導くものと期待される。

さらに、シミュレーション研究では、実際に存在すると仮定されたランダムな項目傾きを無視したとき、特にサンプルサイズが大きく項目数が小さい場合に、Type I error が大幅に上昇することを示した。また、心理学研究で扱う心理尺度を含む実データ分析の多くで、実際に 0 でないランダムな項目傾きの分散が推定され、そのため既存の（ランダムな項目傾きを仮定していない）分析においては偏りのある結論が導かれている可能性を指摘した。加えて、SEM に基づく（実際に存在すると仮定されたランダムな項目傾きを無視した）分析を行った場合、必ずしも適合度指標等は低い水準値を示さないことも確認した。

Rosseel, Y., & Loh, W. W. (2024). A structural after measurement approach to structural equation modeling. *Psychological Methods*, 29(3), 561-588. <https://doi.org/10.1037/met0000503>

報告者：藤本茉里恵

共分散構造分析(SEM)は潜在変数間の関係の強さを調べるため幅広い分野で使われている。しかし、SEMには以下の問題点がある。最尤法(ML)などの反復推定の際に解が収束せずに不適解になること、大きなサンプルサイズを必要とすること、そして測定モデルと構造モデルを同時に推定するために、測定モデルにありうる誤設定の影響が潜在変数間の関係を示す構造モデルに波及することなどである。このため、潜在変数を因子得点や合計得点などの観測変数に置き換えて変数間の相関を調べる方法も多用される。しかし、このような方法では測定モデル内の誤差を含んだまま構造部分が推定されてしまう。

そこで、本論文ではこれらの欠点を乗り越える新たなアプローチ "structural after measurement" (SAM)を提案する。SAMは、測定部分と構造部分を別々に推定するため、仮に測定モデルに誤設定がある場合や、またサンプルサイズが小さい場合においても、より高い精度で潜在変数間の関係を推定できる。

測定部分では、3つの選択肢がある。複数の潜在変数をまとめて一つのCFAモデルとして推定する方法($B = 1$)、各潜在変数に対応するCFAモデルを設定する方法($B = M$)、そしてこれらの中間的な方法($1 < B < M$)である。異なる潜在変数に紐づいている観測変数の残差に相関があるときや、一つの観測変数が複数の潜在変数と関連をもつときは一つのCFAモデル($B = 1$)を用いて推定するのが良いが、異なる潜在変数に紐づいている観測変数同士の残差相関がないときには、 $B = M$ を推奨する。また、非反復推定の利用が推奨されている。

構造部分では、潜在変数の平均ベクトルと分散共分散行列を算出することにより、測定部分から切り離して構造部分を推定するLocal SAMと、固定した測定部分のパラメータと一緒に構造部分を推定するGlobal SAMの2つの推定方法が提案された。Local SAMでは、測定部分と構造部分のそれぞれに適合度を算出できるが、Global SAMでは構造部分に特化した適合度の算出が難しくなる。基本的にはLocal SAMを推奨されるが、Local SAMでは高次因子モデルのように(高次の)潜在変数が他(低次の)潜在変数を説明するモデルへの適用が難しいなどの限界点もあるため、このような場合はGlobal SAMが推奨される。

本論文で紹介されたSAMは、Burt(1976), Hunter and Gerbing (1982), やLance, Cornwell, and Mulaik (1988)により提唱されていた方法を統合したものと位置付けられる。さらに、近年広まりつつある、因子得点を予測してから潜在変数間の回帰を行う方法(factor score regression; Croon, 2002)よりも汎用性が高い方法である。同時推定としてのSEMの問題点を克服する新たな方法として、今後の更なる発展が期待される。

Herderich, A., Freudenthaler, H. H., & Garcia, D. (2024). A computational method to reveal psychological constructs from text data. *Psychological Methods*. Advance online publication.

報告者：土田高太朗

心理学的構成概念は直接観測されないため、正確かつ包括的に定義することが困難である。今日、構成概念の定義には量的なアプローチと質的なアプローチが存在するが、量的なアプローチは質問紙作成などのプロセスにおいて専門家のバイアスを取り込んでしまうなどの可能性があり、一方で質的なアプローチは方法の一般化が難しく、調査者の手技にその結果が依存してしまい、またコストが大きいことサンプルサイズが制限されてしまうなどの問題がある。

本論文では、収集したテキストデータに対して NLP を用いた分析を行うことで構成概念の定義を検討する新たな手法 “Construct Mining Pipeline” を提案する。提案手法は自由記述により得られたテキストデータに対して、NLP を応用して埋め込み表現を取得し、クラスタリングを行うことで、構成概念の構造を明らかにするものであり、これにより定量的なアプローチと質的なアプローチの利点を組み合わせることを試みている。

“Construct Mining Pipeline” は 9 つのステップで構成されている。Step 1: Data Generation においては、対象の心理学的構成概念に関する複数のオープンエンド型の質問を作成し、調査参加者から構成概念に関連する短文を収集する。次に Step 2: Sentence Embeddings で、BERT などの言語モデルを使用することで、収集された文に対応する高次元空間における数値的表現を取得する。そして、Step 3: Item Bias Measurement、および Step 4: Item Bias Reduction で、Step 1 の質問項目による、構成概念に本質的には関わらないバイアスを検出し、軽減を試みる。以上のプロセスで得た、文の高次元空間における数値的表現データに対して、Step 5: Dimension Reduction、Step 6: Clustering で、順に次元縮約とクラスタリングを適用する。次元縮約とクラスタリング、それぞれのステップにおいて、テキストデータと相性がよい方法（UMAP, HDBSCAN）が提案されている。Step 7: Robustness Check において、クラスタリングの結果の安定性を確認し、Step 8: Cluster Validation において、さらに専門性をもたない人々に対して調査を実施することでクラスター内の一貫性を確認する。そのように得たクラスタについて最後に Step 9: Cluster Interpretation で専門家による解釈を行う。解釈の段階では、既知の構成概念を参考にすることや、ワードクラウドなどにより出現頻度が高い単語を抽出することが有効である。

本論文では提案手法を用いて感情調整（emotion regulation）の構成概念の構造を検討しており、その結果、15 の感情調整のクラスタが抽出された。その一部は先行研究によって言及されているものが含まれている一方で、未だ注目されていないものも含まれていた。

Kroc, E., & Olvera Astivia, O. L. (2023). The case for the curve: Parametric regression with second- and third-order polynomial functions of predictors should be routine. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000629>.

報告者：佐々木一洋

多項回帰は古くから広く利用されている手法であり、また非線形モデリングの方法として初学者への教授に適したものである。しかし、教科書によっては使用を避けるように促すなど、その重要性が十分に認知されているとは言えない。本研究ではまず、多項回帰の基本的な意義を確認し、そのうえで24の北米で用いられている教科書において多項回帰がどのように扱われているかを概観する。その後、多項回帰がどのような場面で頑健で有用であるかを説明し、多項式回帰についての誤解を紐解いていく。

多項回帰は変数間の曲線的な関係を捉えるうえで有用とされてきた。にもかかわらず、調査の対象とした24の統計学ないし回帰分析の教科書の半数以上において多項回帰が十分に扱われておらず、また使わないことを明示的に推奨する記述が含まれる教科書もあった。これらの教科書では、数学的知識が不十分な初学者には概念的理解のコストが大きくまた数的処理もブラックボックスとなってしまうやすいノンパラメトリックの手法が推奨されていた。

これに対し多項回帰は、特に微積分に十分に習熟していない初学者であっても線形回帰分析の概念を少し応用することで、新しい内容を多く学ぶ必要もなしに理解することができる。局所線形性の考え方を導入することで、微積分を前提とせずに、変数間の大域的関係と局所的関係とを捉えることが可能になる。こうした点で、多項回帰は幾つかのノンパラメトリックの手法と比べて初学者への導入に適している。

また、実際の例を基に、二次や三次の項を含めることで天井効果や床効果を効果的にモデリングできることを例示した。特に、傾きの正負の変化が一度のみのグラフにおいては従来二次のモデルが有用とされてきたが、三次の項も含めることで変化率の変化を捉えられる。

こうした利点に対し、多項回帰はいくつかのいわれなき批判を受けている。解釈の難しさ、計算的不安定性、共線性の問題、過適合などの批判は、多項回帰に限った問題でなく、また機器の改良や、情報量規準の利用および十分なサンプルサイズの確保などを通して対処可能な問題である。唯一、コンパクトな変数でない場合にあってはまりが悪くなるという、漸近的なふるまいを制御できない特性には教授上の注意が必要である。しかし、この点を踏まえても、初学者が意味を理解して用いるうえで、多項回帰は曲線的な関係をさぐるためのルーチンとして標準的に教授される内容となっていくべきである。

Sijtsma, K., Ellis, J.K. & Borsboom, D. (2024). Recognize the Value of the Sum Score, Psychometrics' Greatest Accomplishment. *Psychometrika*, 89, 84-117.

報告者：中島優稀

心理学的なテストにおける「合計点」は前科学的な概念としてみなされることが多く、「潜在変数」に基づく測定法に代替されるべきだという意見が増えている。このような意見は、合計点による測定は潜在変数モデルによる測定の簡易版に過ぎないという考えや、古典的テスト理論（CTT）によって合計点の信頼性の推定が可能になるのは一次元性があるときのみで、現実的な状況に適合することは稀だという考えに基づいている。本稿は、現代の測定モデルには CTT では達成できない独自の長所があることを認めつつも、CTT はその一般性と実用性によって現代の測定モデルを補完できることを主張するものである。

第一に、項目反応理論（IRT）の文脈において、離散的な合計点は連続的な潜在変数の順序づけの指標として捉えられる点である。すなわち、潜在変数を推定できない場合でも、合計点で順序づけることによって潜在変数の平均値で順序づけられるということである。この特性は潜在変数の確率的順序（SOL）と呼ばれる。さらに、項目反応関数が完全には特定されていなかったりサンプルサイズが小さかったりするような現実的な状況では、合計点は重みつき最尤推定量よりも潜在変数の推定値として優れていることがシミュレーションにより示された。また、項目が共通の潜在変数に依存するのではなく、項目どうしが相互作用するような場合において、ネットワークモデルが提案されている。こうしたモデルで合計点を位置づけるために、統計力学で用いられるイジングモデルを検討したところ、ネットワークの期待される全体的な状態、すなわち真のネットワークを評価する上で、合計点が有用であることがシミュレーションにより示唆された。

第二に、観測得点を真の得点と誤差得点に分解する CTT のモデルの下では、 α 係数は観測得点の信頼性の下限値として捉えられる点である。信頼性は観測不可能な真の得点の分散に依存するが、 α 係数は観測可能な変数の関数で表されるため、信頼性について重要な情報を提供する。特に重要なのは、CTT の仮定は誤差得点と真の得点の無相関性および誤差得点どうしの無相関性のみである点であり、このような CTT の仮定の緩さによって心理測定における多大な貢献が期待されている。信頼性が α 係数に一致する場合には一次元性が満たされるが、一次元性がないと α 係数は有用でないというのは誤解であり、むしろ一次元性がないような現実的な状況でも信頼性のとりうる値の範囲を特定できることにこそ価値がある。CTT は信頼性の評価において有用であるため、測定モデルというより誤差モデルとして理解される。ゆえに、真の得点が必ずしも心理的特性と一致しないような場合でも役に立つため、CTT の適用範囲はあらゆる文脈に広がりつつある。

Loh, W. W., & Ren, D. (2023). Estimating time-varying treatment effects in longitudinal studies. *Psychological methods*, Advance online publication. <https://doi.org/10.1037/met0000574>

報告者：加藤 剛

本論文では、縦断研究における時間依存性交絡を伴う処置効果の推定手法として、線形構造ネスト平均モデル (SNMM) に基づく g 推定 (g-estimation) を紹介する。従来、時間とともに変化し、同時に処置によって影響を受ける時間依存性共変量 (time-varying covariate) を標準的な回帰分析で調整するか除外するかのいずれかでは、どちらの方法でもバイアスが生じ得るという問題があった。本研究では、時間依存性共変量を適切に取り扱いながら、複数の時点にわたる処置効果を推定できる g 推定量の有用性を示す。

g 推定は、二重に頑健 (doubly robust) な推定法であるという点に大きな特徴がある。すなわち、傾向スコアモデル (処置モデル) とアウトカムモデルのうち、少なくとも一方が正しく特定されていれば、推定量は一致性をもつ。また、時間不変共変量・時間依存性共変量いずれの場合も、g 推定ではその分布のモデル化を必要としないため、柔軟かつ実装が比較的容易である。さらに、連続変数・非連続変数どちらの処置にも対応可能であり、効果修飾因子 (例：年齢や不安など) との交互作用を加味することで、効果の異質性を詳しく検討することができる。

本論文では、まず単一時点の処置効果推定における g 推定の手順を示し、続いて時間依存性処置へ拡張する方法を説明する。時点 $t = 0$ から $T - 1$ までの各処置に対して傾向スコアモデルを構築し、残差化アウトカムを段階的に更新することで、各時点の処置効果を求められる。また、時間依存性アウトカムを含む状況にも SNMM を適用することで、たとえば途中の時点で測定されたアウトカムに対しても同様の推定を実施できる。

シミュレーション研究により、g 推定量はモデルの一部を誤特定していてもバイアスのない推定が可能であることを経験的に示し、他の単純な推定量に比べて優位性をもつことを確認した。さらに、実例としてメディア使用が生活満足度に与える影響を扱った縦断データを用い、g 推定を適用した結果を示した。そこでは、書籍の使用による平均的な影響はほとんど見られないことが推定される一方、年齢や不安などの効果修飾因子に着目することで、時点により変化する因果効果の違いを考察できることが示唆された。

g 推定は未測定交絡がないという仮定の下で、時間依存性共変量を含む複雑な縦断データにおいても、処置効果を二重に頑健な形で推定可能にする手法である。心理学における縦断研究で時間依存性処置の因果推論を行う際には、g 推定の使用が推奨される。

Jiménez, M., Abad, F. J., Garcia-Garzon, E., Golino, H., Christensen, A. P., & Garrido, L. E. (2023). Dimensionality assessment in bifactor structures with multiple general factors: A network psychometrics approach. *Psychological Methods*. <https://doi.org/10.1037/met0000590>

報告者：吉田 欧太

本研究は、各項目が一般因子とグループ因子に同時に負荷する 2 因子モデル（バイファクターモデル）における因子数の評価方法について分析するものである。1 つまたは複数の一般因子を持つ構造に対する因子保持法の精度は、こうした構造が、知能、パーソナリティ、精神病理学といった分野で一般的に見られるにも関わらず、次元に関する研究において見過ごされがちであった。本研究では、この問題を解決するために、いくつかの因子保持法の性能を比較した。

本研究では、まずグループ因子数を推定するために、カイザー基準、経験カイザー基準、主成分や主因子法による平行分析（前者を PA_{PCA} ）および Louvain アルゴリズムを活用したクラスタリングによる探索的グラフ分析（ EGA_{LV} ）を因子保持法として対象とした。本研究で開発された EGA_{LV} は、ネットワークサイコメトリクスのアプローチを活用し、GLASSO 正則化を伴ってガウシアングラフィカルモデル（GGM）を適用することによって変数間の偏相関を推定する EGA に、Louvain のアルゴリズムを活用してクラスタリングを行う。

次に、この中で最良であった PA_{PCA} と EGA_{LV} という 2 つの方法によって提案された 1 次解の因子スコアを用いて、一般因子の数を推定し、 PA_{PCA} と EGA_{LV} の 2 次バージョンを得た（それぞれ PA_{PCA-FS} 、 EGA_{LV-FS} ）。これらは、Goldberg の Bass-Ackwards 法と異なり、低次の因子の数を推定し、次に高次の因子を推定するという意味で、ボトムアップアプローチである。さらに、 EGA_{LV} が提供するマルチレベルの直接解も検証した。

すべての手法を、母集団との誤差を含む 9 つの変数を操作した大規模シミュレーションで評価した。その結果、 PA_{PCA} と EGA_{LV} は、真のグループ因子数を探索する上で、総合的に最も優れた性能を示したが、前者は弱いグループ因子と小さなサンプルサイズに対して、後者は高いクロスローディングに対して、より影響を受けやすいという結果であった。一般因子数の推定に関しては、 PA_{PCA-FS} と EGA_{LV-FS} は、すべての条件において完璧に近い精度を示したが、 EGA_{LV} は不正確であった。EGA に基づく手法は、実際に遭遇する可能性の高い条件に対して頑健であった。その後、両者を実データ（The HEXACO-100 Inventory）に適用したが、サンプルサイズやグループ因子ごとの指標の数が少なかったこともあり、 PA_{PCA} は因子数を過小評価した一方、 EGA_{LV} は理論モデルに適合した結果を得ることができた。