

再現性問題を踏まえた 構造方程式モデリングの 研究応用のあり方

お茶の水女子大学 伊藤大幸

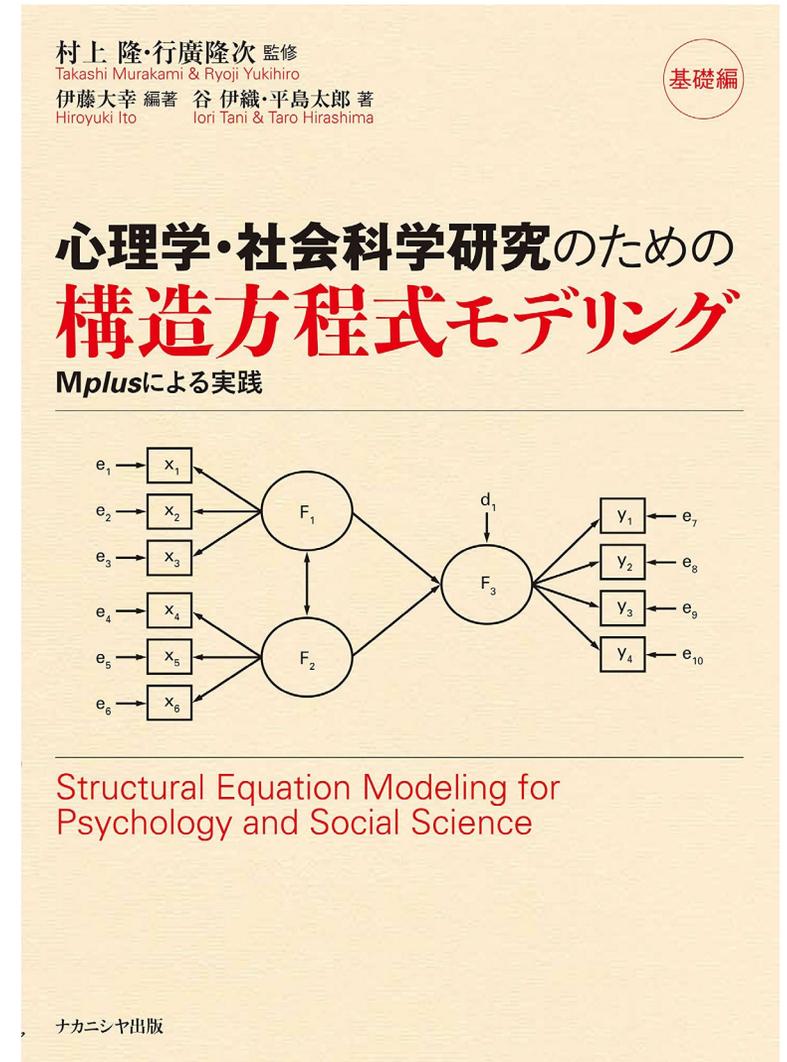
自己紹介

◆ 専門

- 心理学全般（認知、発達、教育など）、研究法・統計法（研究応用）

◆ 研究

- 認知と感情の相互作用
- 子どもの発達とメンタルヘルスに関する大規模縦断研究
- 発達障害や関連特性のアセスメント
- 多変量解析の研究応用



構成

- ◆ 序論：SEMの現状と再現性問題
- ◆ 本論：SEMの教育
- ◆ 結論

構成

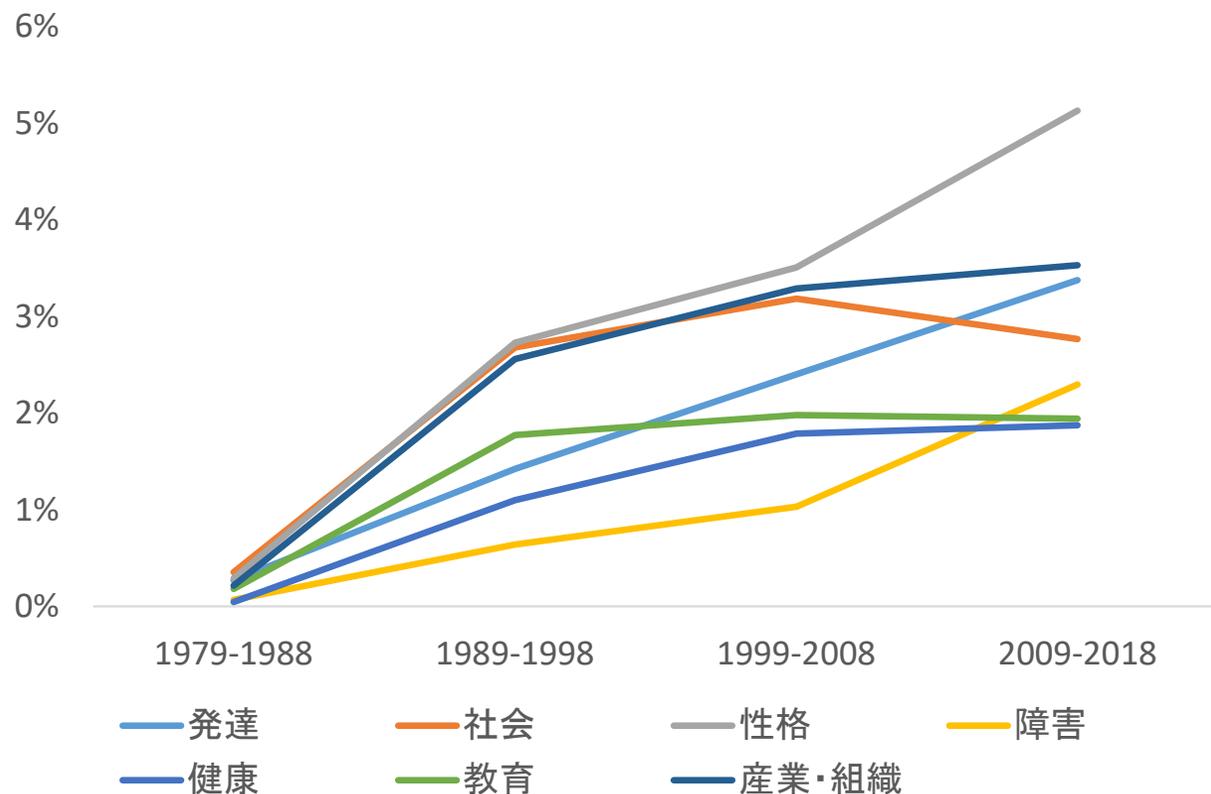
- ◆ 序論：SEMの現状と再現性問題
- ◆ 本論：SEMの教育
- ◆ 結論

SEMの柔軟性・包括性

- ◆ 社会科学領域で用いられる解析手法の大部分を下位モデルとして扱える
 - ✓ 一般線形モデル（t検定、分散分析、重回帰分析など）
 - ✓ 一般化線形モデル（ロジスティック回帰分析など）
 - ✓ 因子分析（探索的、確認的、階層的など）
 - ✓ マルチレベルモデル
 - ✓ 項目反応理論 など
- SEMを正しく理解することは、現在の心理統計の技術で何ができて、何ができないかを知ることと、ほぼ同義

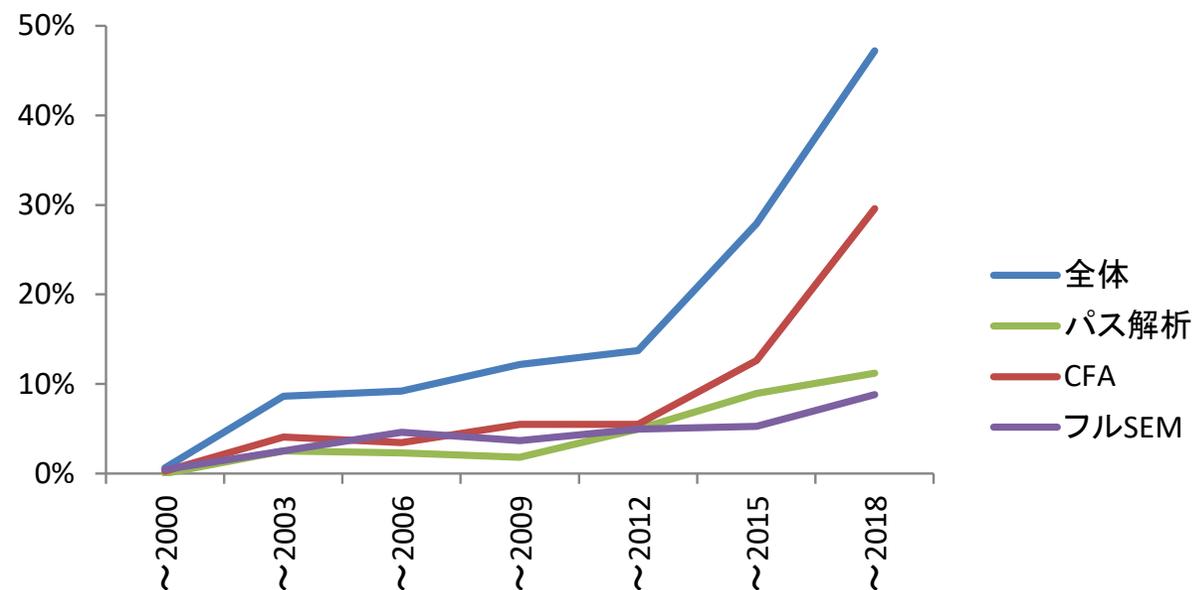
SEMの利用動向

世界



※ 『PsychINFO』においてキーワードを「"path analysis" or "confirmatory factor" or "latent variable" or "structural equation"」として検索

国内



※ 『心理学研究』掲載論文に占める割合

SEMは心理学の発展に貢献しているか？

- ◆ Big numbers, low quality (Kline, 2023)
 - SEMの「人気」
 - ✓ 便利なコンピュータプログラムの普及
 - ✓ 扱えるリサーチクエスチョンの広さ
 - SEMの応用研究の質の低さ
 - ✓ SEMを使用した大部分の研究が何らかの深刻な欠陥を有する

SEMの不適切な使用・報告

◆ Zhang et al. (2021)

- 組織心理学のトップ12誌をレビュー
 - ✓ SEMを使用する理由の説明がない：58%
 - ✓ 分布の仮定が検証されていない：82%
 - ✓ モデルの詳細が示されていない：24%
 - ✓ 欠測値への対処が示されていない：44%
 - ✓ 測定モデルと構造モデルが区別されていない：19%
 - ✓ 部分的適合の情報が報告されていない：83%

SEMの不適切な使用・報告

◆ 伊藤（2019）

- 2015～2018年度に掲載された「心理学研究」の掲載論文
 - ✓ 因果関係の根拠が示されていない：30%
 - ✓ 項目・概念の重なり：60%
 - ✓ 測定モデルと構造モデルが区別されていない：90%
 - ✓ 不明瞭なモデル指定：33%
 - ✓ 不自然な結果：17%
 - ✓ 相関行列が報告されていない：60%
 - ✓ 適合度指標の恣意的な報告：50%
 - ✓ 因果関係の不適切な解釈：30%

SEMと再現性

◆ Kaplan (2009)

- モデル適合度や修正指標に基づく事後的なモデル修正が横行し、研究知見の再現性が損なわれている
- SEMは政策や臨床に関する重要な問題の予測研究にはほとんど使用されていない

◆ Kline (2023)

- SEMの研究のほとんどは追試されていない単発研究
- SEMは**再現性の危機**の一部を構成している

再現性の危機

◆ Why Most Published Research Findings Are False? (Ioannidis, 2005)

- 再現性を低める要因
 1. 小さいサンプルサイズ
 2. 小さい効果量
 3. 探索的な研究
 4. デザインや分析手法の**柔軟性**
 5. 利害や偏見の大きさ
 6. 研究テーマの流行

SEMの柔軟性は
短所にもなりうる

再現性を低める研究行為

- ◆ 疑わしい研究行為（Questionable Research Practices: **QRP**）
 - **HARKing**：結果に合わせて仮説を後付けする
 - **Cherry-Picking**：多くの分析結果の中から都合のよい結果だけを選択的に報告する
 - **p-hacking**：有意な結果が得られるようデータや分析方法を操作する

SEMの（開示されない）柔軟性は全てのQRPを助長しうる

序論の結論：SEMとどう付き合っていくのか

- ◆ SEMは統計学の英知の結晶
 - 道具自体に罪はない
- ◆ しかし、現状では心理学の発展に貢献するどころか、発展を妨げている
 - 開発者（統計学者）はあらゆる局面における道具の使い方を懇切丁寧にサポートしてくれるわけではない
 - ユーザーと開発者の対話を通して、ベストプラクティスを見出していく必要がある

構成

- ◆ 序論：SEMの現状と再現性問題
- ◆ 本論：SEMの教育
- ◆ 結論

SEMの教育

- ◆ 学部や大学院のカリキュラムでSEMの系統的な教育を実施している大学は多くないのでは？
 - 一からしっかり教えるには1クール分程度の回数が必要
 - ✓ SEMだけにそれだけの回数を設定することは難しい
 - 学生や院生が先行研究やテキストを読んで、独学で学んでいるケースが多い？
 - ✓ 本質を理解せずに形だけを真似する形になりやすい（それが比較的容易にできてしまう）

SEMの教育

- ◆ 学生（あるいは心理学研究者）に最低限伝えたいこと
 1. SEMを使用する必要性の吟味
 2. 構造モデルと測定モデルの区別
 3. 理論・仮説の重要性
 4. 交絡因子の統制
 5. 適合度の意味
 6. 透明性のある報告

1. SEMを使用する必要性の吟味

◆ 重回帰分析で事足りるのでは？

- SEMが有効となる状況（パス解析の文脈で）
 - ✓ 媒介関係を含む複雑な因果モデルが想定される
 - ✓ 潜在変数を用いて測定誤差に対処したい（消極的な潜在変数利用）
 - ✓ 潜在成長モデル、ランダム切片交差遅延モデルなど、潜在変数を積極的に利用したモデルを使用する
 - ✓ 双方向の因果関係や因果ループを含む非再帰的モデルを扱う
- 重回帰分析で事足りる状況
 - ✓ 内生変数（従属変数）が一つしか想定されない
 - ✓ 内生変数間の因果関係が仮定されない（または仮定できない）

横断データに複雑なモデルをあてはめることの妥当性も吟味

1. SEMを使用する必要性の吟味

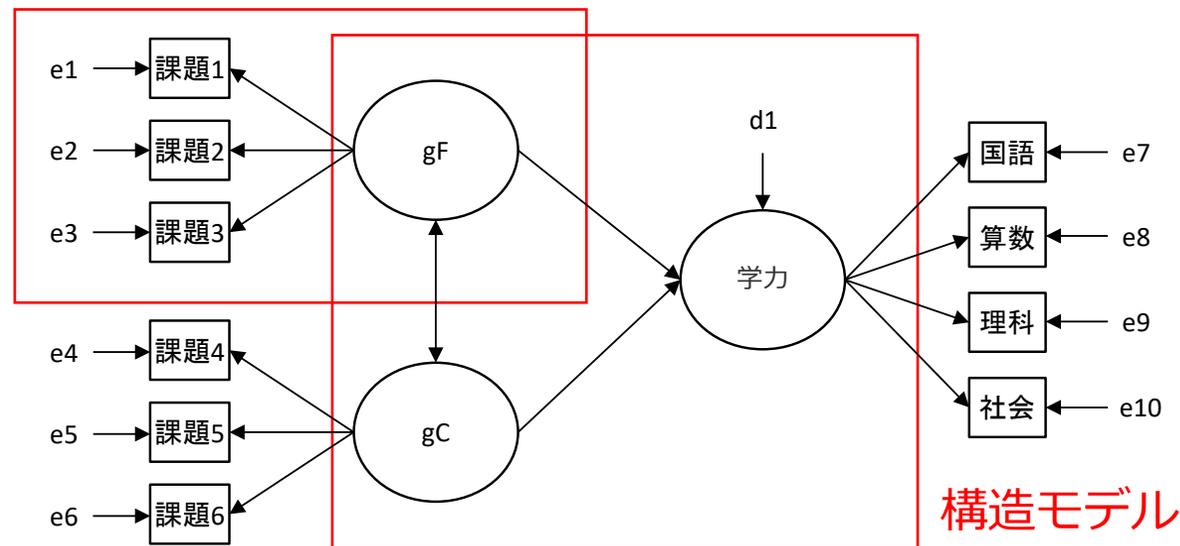
- ◆ 探索的因子分析の方が適切なのでは？
 - SEMが有効となる状況（確認的因子分析の文脈で）
 - ✓ 先行研究で因子構造が確認されている
 - ✓ 複数の因子モデルを比較したい
 - ✓ 複数の集団間の測定不変性を検証したい
 - ✓ 共通の方法因子を考慮したい（逆転項目、評定者など）
 - 探索的因子分析に適した状況
 - ✓ 因子構造に関する仮説やそれを裏付ける先行研究の知見がない
 - ✓ 因子構造に関する仮説はあるが、尺度が独自に作成され、比較対象となる対立仮説もない

2. 構造モデルと測定モデルの区別

◆ SEM = 確認的因子分析 + パス解析

- 確認的因子分析（**測定モデル**）：観測変数の背後にある構成概念（潜在変数・因子）に関する仮説を検証
- パス解析（**構造モデル**）：構成概念間の因果関係に関する仮説を検証

測定モデル



構造モデル

2つのモデルを区別せず扱うことが大部分の不適切な運用の根本的原因

2. 構造モデルと測定モデルの区別

◆ (例えば) 誤差相関の扱い

- 測定モデルの場合

- ✓ 観測変数間の関連を潜在変数（因子）によって説明することに目的があるため、因子で説明できない誤差相関を（実質的な理由なしに）想定することは、モデルの欠陥を意味する

- 構造モデルの場合

- ✓ 独立変数と従属変数の関連を検証することが目的であり、従属変数間の関連を独立変数が全て説明することを想定するわけではないので、（実質的な理由がなければ）従属変数間に誤差相関を想定することが合理的

実際は誤差相関を設定していない誤用が多い

3. 理論・仮説の重要性

- ◆ SEMは先験的な理論・仮説の検証を目的とした**確証的**な分析の枠組み
 - 理論・仮説が明確でない段階での**探索的**な分析には適していない
- ◆ 測定モデルの場合
 - 修正指標などに基づく過度なモデル修正はデータへの**過剰適合** (overfitting) により再現性の低い結果をもたらすリスク
 - 同一のデータで探索的因子分析→確認的因子分析はNG

3. 理論・仮説の重要性

◆ 構造モデルの場合

- 因果関係の方向は仮定されるものであって、検証されるものではない
 - ✓ 方向の仮定に基づいて、強度を推定するのがSEMの役割
 - 心理学で扱う構成概念について、因果関係の方向を先験的に論証することは容易でない
 - ✓ 心理学理論の未成熟
 - ✓ 双方向の因果関係
- 横断データで複雑な因果モデルを設定してパス解析を行うことは適切でない場合が多い

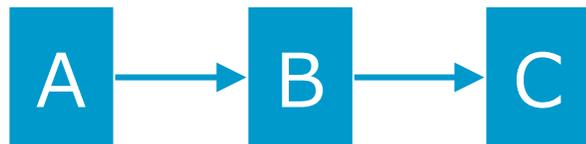
3. 理論・仮説の重要性

◆ 同値モデルの考慮

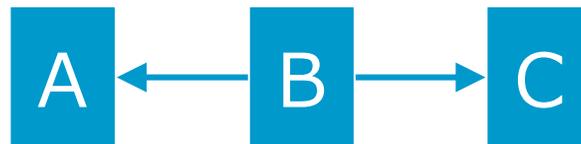
- 仮説モデルと等しい適合度を与えるモデル
- Lee & Hershberger (1990)の置換ルール
 - ✓ 他の変数からのパスを受けない変数ブロックが**飽和モデル**（全変数間にパスが引かれたモデル）であるとき、そのブロック内のどの関連も、異なる方向や種類の関連に置き換えることができる

因果関係の方向を適合度で検証することは基本的に困難

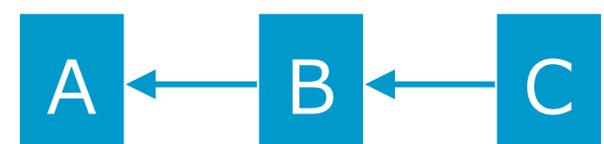
仮説モデル



同値モデル 1



同値モデル 2



4. 交絡因子の統制（構造モデル）

◆ 因果関係を立証するための条件

1. 仮説上の原因と結果に相関がある
 2. 仮説上の原因が結果よりも時間的に先行している
 3. 1の関連が第三の変数（**交絡因子**）によって説明されない
- 全ての交絡因子を統制しなければ因果関係を正確に推定できない

4. 交絡因子の統制（構造モデル）

- ◆ 心理学領域では、医学、経済学などに比べ、交絡因子と考えられる変数が統制されていない研究が目立つ
 - 国内では年齢、性別などの基本的属性さえ統制されていない研究も多数
- ◆ 重回帰分析に関する吉田・村井（2021）の指摘
 - 交絡変数は組み入れる必要がある
 - 1つの構成概念に関する複数の観測変数を同時に組み入れることは適切でない

相互に独立した概念に関する変数も同時に組み入れてはいけないという誤解が生じている

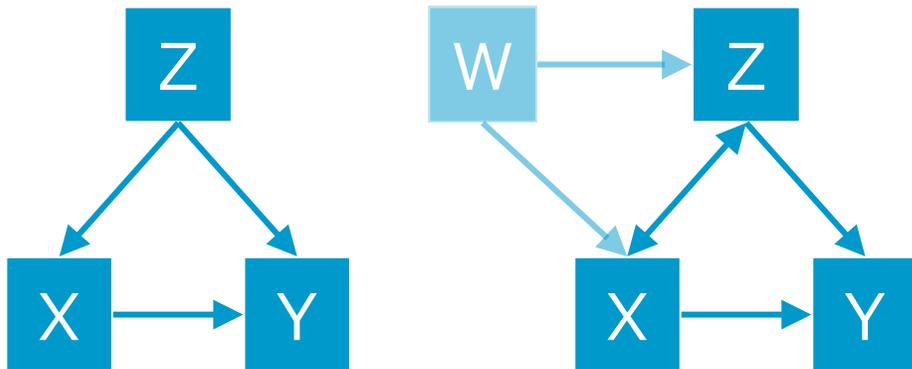
4. 交絡因子の統制（構造モデル）

◆ グラフ理論（Pearl, 2009）

• 統制すべき変数

- ✓ 共通の原因変数（交絡因子）
- ✓ 共通の原因変数により相関が生じている変数

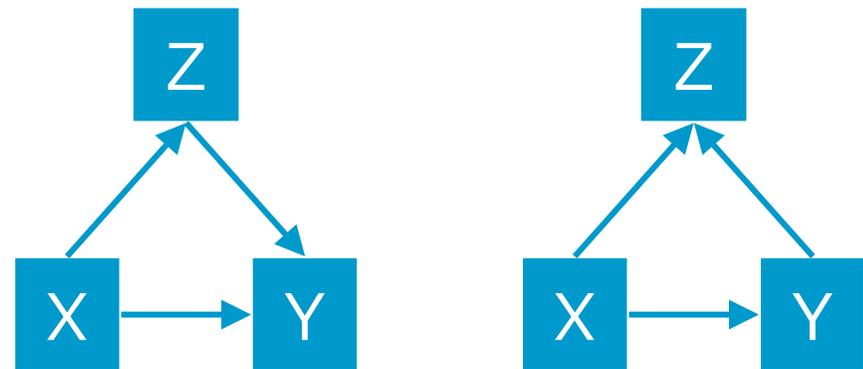
独立変数から見て
上流側にある変数や
並列位置にある変数



• 統制すべきでない変数

- ✓ 媒介変数
- ✓ 共通の結果変数（合流点）

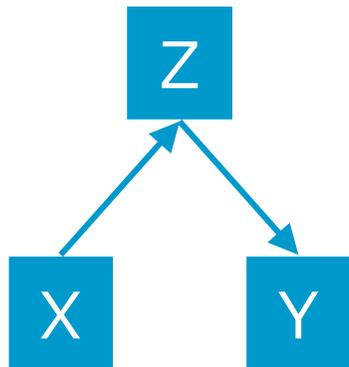
独立変数から見て
下流側にある変数



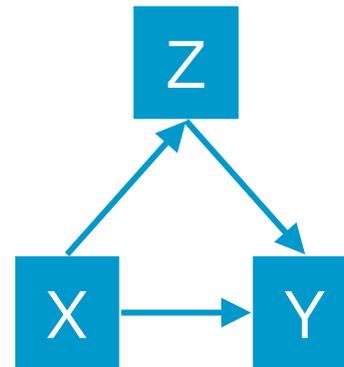
5. 適合度の意味

◆ 構造モデルの場合

- そもそも適合度を議論することに意味がない場合が多い
 - ✓ 適合度はモデルで想定されていない関連が実際にはない（小さい）ことを示唆するにすぎない
 - ✓ 因果関係の方向性を支持するものではない（同値モデルの問題）
 - ✓ 因果関係の強度を示すものでもない（パス係数で示される）



適合度の高さは $X \rightarrow Y$ の直接効果が小さいことを意味するだけ



飽和モデルで $X \rightarrow Y$ のパス係数を推定した方がわかりやすい

5. 適合度の意味

◆ 構造モデルの場合（つづき）

- 適合度指標のカットオフ値（CFI>.95、RMSEA<.05、SRMR<.08など）は測定モデルのシミュレーションで見出された値であり、構造モデルに適用するのは不適切

→ 構造モデルの適合度はカイ二乗値によって評価するのが原則

- ✓ サンプルサイズによらず、カイ二乗値が有意であれば、本来有意になるはずのパスが設定されていない可能性を強く示唆するので、モデル修正の必要がある

「サンプルサイズが大きいとカイ二乗検定は有用でない」
は測定モデルの話！

5. 適合度の意味

◆ 測定モデルの場合

- 固定的な適合度指標のカットオフ値（CFI>.95、RMSEA<.05、SRMR<.08など）は特定の状況でのシミュレーションに基づく目安にすぎない
 - ✓ 因子あたりの項目数、因子負荷量、因子間相関などの要因によって、適合度指標の分布は大きく変動する
- 特定のモデルとデータセットに応じて調整された**動的適合度基準**（Niemand & Mai, 2018）を使用するのが望ましい
 - ✓ 実際のモデルとデータに基づくシミュレーションにより算出

5. 適合度の意味

- ◆ 構造モデルと測定モデルの両方を含むモデルの場合、両者の適合を分けて評価（**二段階評価**）
 1. 全ての潜在変数（因子）間に相関を仮定した確認的因子分析モデルの適合を評価
 - ✓ 適合に問題がなければ2に進む
 2. 潜在変数間に当初想定した構造モデルをあてはめ、1からのカイ二乗値の変化を検証
 - ✓ カイ二乗値の差が有意でなければモデルが支持される

6. 透明性のある報告

◆ 指定したモデルの明示

- 主要な部分のパス図だけでなく誤差相関や等値制約などを含む完全な詳細を記述
- 可能な限り、プログラムのコードを示す
- **事前登録**によりあらかじめモデルを明示することが最善

◆ 修正過程の明示

- どのような根拠（実証的および理論的）に基づいて、どのような過程で修正を施したかを詳細に記述

6. 透明性のある報告

◆ 相関行列（分散共分散行列）の報告

- SEMの推定結果はローデータがなくても、（基本的に）観測変数間の相関行列があれば再現可能
 - ✓ 変数が多い場合もサプリメントに示す
- 他の研究者が追試だけでなく、異なるモデルを検証してみることも可能に
- メタ分析にも利用できる

構成

- ◆ 序論：SEMの現状と再現性問題
- ◆ 本論：SEMの教育
- ◆ 結論

結論

- ◆ SEMが必要となる状況は限られる
- ◆ 因子分析とパス解析は別物
- ◆ 理論がなければSEMは使えない
- ◆ 考えうる交絡因子を全て測定し、統制する
- ◆ 適合度指標の固定的なカットオフ値は使わない
- ◆ モデルと結果の完全な詳細を報告