

マルチレベル分析について

清水裕士

広島大学 大学院総合科学研究科

simizu706@hiroshima-u.ac.jp

<http://norimune.net>

はじめに

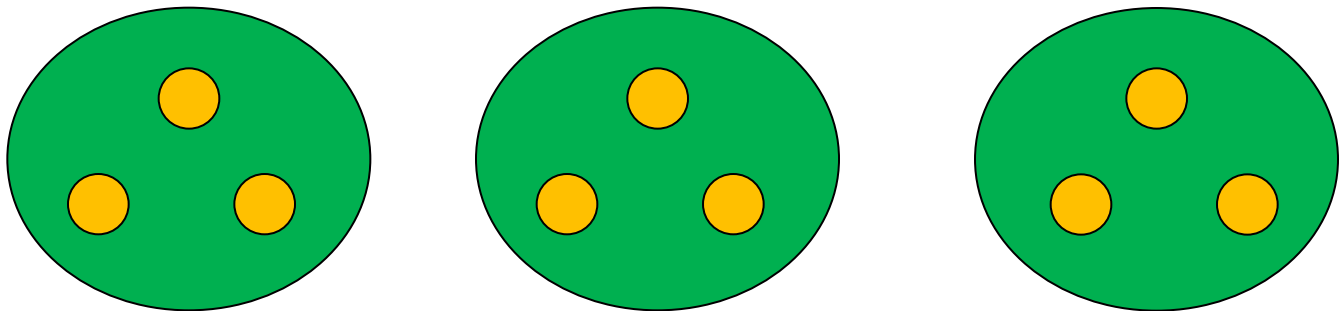
- マルチレベルモデルの2種類のタイプ
 - 継時データに対する分析
 - 同じ患者さんから継続的にデータを取るような場合
 - →潜在曲線モデルや階層線形モデル(HLM)が得意
 - 集団にネストされたデータに対する分析
 - 病院や病棟単位で得られたデータ
 - →マルチレベル構造方程式モデル(ML-SEM)が得意

本発表の概要

- なぜマルチレベルモデルを使うのか
 - 階層的データの特徴と従来法の問題点
 - マルチレベルモデルの特徴
- マルチレベル構造方程式モデルについて
 - ML-SEMの実践
 - SEMとの違い, 実際に分析した感じ
 - ML-SEMの解釈の仕方
 - 結果の見方, 論文の書き方

個人-集団データの階層性

- データの階層性
 - 集団ごとにネストされたデータ



- 集団ごとに共通した値が入力されるデータ
- 集団内で類似したデータ
 - 学校-生徒、カップルデータ、反復測定データ・・・etc

このようなデータを階層的データと呼ぶ

サンプルデータ紹介

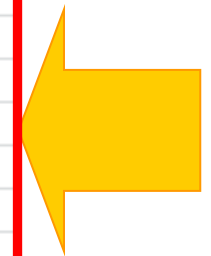
- 仮想的なデータを利用
 - 3人集団が集団討議を行う実験(100集団300人)
 - ※実際に実験は行っていません!
 - 何が課題満足度を高める要因となるのか?
- 測定変数
 - 発言量 → 録音してコーディング
 - 課題の満足度 → 実験後測定
 - 集団パフォーマンス → 集団単位で採点
 - コミュニケーションスキル → 実験前に測定
 - 実験条件 → 0が同じ情報条件、1が違う情報条件

階層的データ

例：サンプルデータの場合

B	C	D	E	F	G
グループ	満足度	発話量	スキル	集団成績	条件
1	3	3	1	3	1
1	3	2	3	3	1
1	3	3	1	3	1
2	3	3	3	3	0
2	2	1	2	3	0
2	2	1	1	3	0
3	1	3	3	2	0
3	3	3	2	2	0
3	3	2	2	2	0
4	1	4	2	1	1
4	1	4	3	1	1
4	1	4	3	1	1
5	2	2	2	5	1
5	3	4	3	5	1
5	2	3	1	5	1
6	3	3	3	5	0
6	3	2		5	0
6	3	2	3	5	0
7	4	3	3	7	1

集団で類似したデータ



集団で一致したデータ

集団単位で平均したデータ

A	B	C	D	E	F	
グループ	満足度	発話量	スキル	集団成績	条件	
1	3	2.666667	1.666667	3	1	
2	2.333333	1.666667	2	3	0	
3	2.333333	2.666667	2.333333	2	0	
4	1	4	2.666667	1	1	
5	2.333333	3	2	5	1	
6	3	2.333333	3	5	0	
7	3.666667	3.666667	2.333333	7	1	
8	3	3	1.333333	3	0	
9	3	2.666667	2.333333	4	0	
10	5	3.666667	2.333333	5	1	
11	3.666667	3.666667	1.666667	8	0	
12	3.333333	3.666667	2	6	1	
13	3.333333	2	1.333333	4	0	
14	3.666667	3	1.666667	6	0	
15	3.333333	2.333333	2.333333	5	0	

従来の方法の問題

- 「サンプルの独立性仮定」の違反
 - 統計学は、サンプルが独立していることを仮定
 - 階層的データは、サンプルが独立していない
 - データの情報量を多く見積もっている
 - つまり、標準誤差を過剰に小さく見積もってしまう
 - →タイプ I エラーを犯す危険がある
- 平均値は集団の性質を反映しない
 - 平均値は集団の性質と個人の性質が混在
 - 得られた相関係数が何を表しているか不明
 - 人数が少ないときほど、危険

普通の相関係数

- N=300の結果
 - 相関係数の有意性検定結果は誤り
 - 相関係数の値も、個人・集団レベルが混在

相関分析								
		満足度	発話量	スキル	集団成績	条件		
	満足度	1.000						
	発話量	.307 **	1.000					
	スキル	.209 **	.141 *	1.000				
	集団成績	.297 **	.075	.053	1.000			
	条件	.188 **	.384 **	.048	-.097 +	1.000		
		** $p < .01$, * $p < .05$, + $p < .10$						

集団単位で平均した相関係数

- N=100の結果
 - 有意性検定は過小評価してしまう
 - 相関係数の値も、個人レベルが依然混在

相関分析		満足度	発話量	スキル	集団成績	条件
満足度		1.000				
発話量		.371 **	1.000			
スキル		.118	.177 +	1.000		
集団成績		.394 **	.101	.095	1.000	
条件		.249 *	.522 **	.079	-.097	1.000
		** $p < .01$, * $p < .05$, + $p < .10$				

相関係数の算出における注意点

- 階層的データの場合の相関係数
 - サンプルサイズが300で検定することの問題
 - データが類似しているのに独立性を仮定すると、推定精度を高く見積もりすぎてしまう → Type I エラー
 - 個人間変動としてのみ解釈することの問題
 - 課題中によく発言した人ほど、会話に満足？
 - それだけの効果ではない可能性
 - つまり、課題中によく話す集団は、みんな満足していた可能性もある(集団間変動としての解釈)

グループ内類似性

- 個人-集団データのグループ内類似性
 - 発言量や満足度は、個々の集団内で類似する
 - →盛り上がっている集団は全員の発話量が多い
- グループ内類似性が階層的データの特徴
 - 個人の得点同士に類似性が見られることによって、サンプルの独立性が違反される
 - 類似性を適切に扱えば、問題は回避される

級内相関係数

- グループ内の類似性を評価するための指標
 - -1~1の間を取る
 - 有意性検定の結果を確認
 - .10以上あれば、マルチレベルモデルを行ったほうがよい、という説もある。

変数名	有効N	級内相関	df1	df2	F値	p値
満足度	300	.358	99	200	2.671	.000
発話量	300	.316	99	200	2.388	.000
スキル	299	-.020	99	199	.941	.630
集団成績	300	1.000	99	200	---	.000
条件	300	1.000	99	200	---	.000

級内相関係数の目安

- 級内相関係数とマルチレベルモデル
 - 級内相関が0より大きくても, とても小さい場合
 - 標準誤差が大きくなる・・・推定が不安定
 - 相関が1を超える・・・不適解
- どれくらいの級内相関があれば, 階層的データ?
 - 明確な基準はないが・・・
 - 集団内人数に関わらず, 0.10以上あれば使ったほうがよい
 - デザインエフェクトが2を超えたら, 使ったほうがよい
 - デザインエフェクト = $1 + (\text{集団内人数} - 1) \times \text{級内相関係数}$

階層的データのまとめ

- 集団ごとにネストされている, 非独立なデータ
 - その本質は, 級内相関が存在するデータ
 - ネストされていればすべてがマルチレベルモデルの対象になるわけではない
- 従来の方法では, 誤った結果・解釈を得てしまう
 - 推定精度をよく見積もりすぎる → Type I エラー
 - 得られた相関係数は, 集団単位・個人単位の効果が混在する
- そこで, マルチレベルモデルが必要

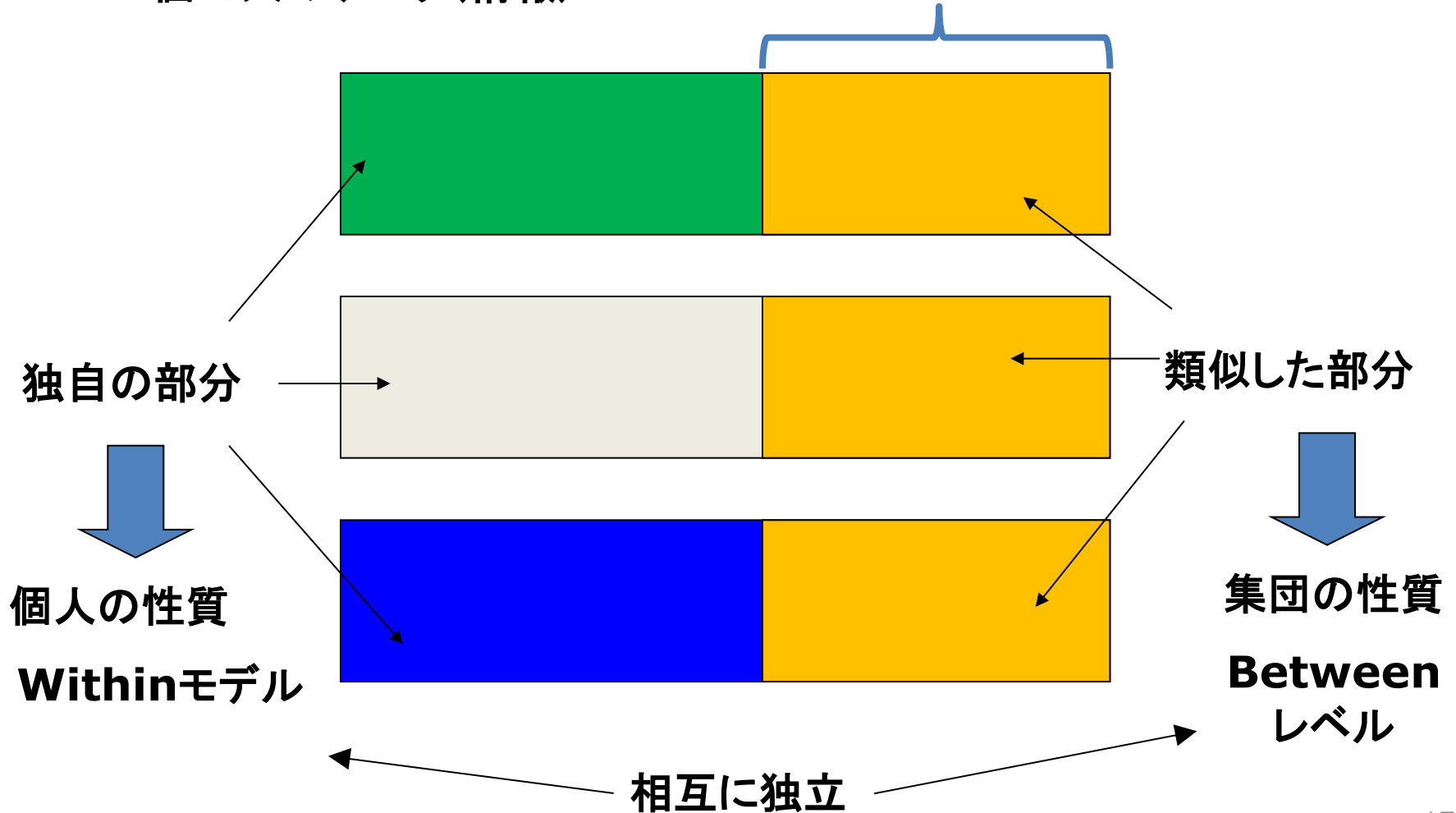
マルチレベル分析

- マルチレベル分析とは？
 - 階層的データを適切に分析する手法
 - 個人・集団だけに限らない
 - 「グループ内類似性」=級内相関係数を評価し、それにあわせたモデリングを行う
- いろいろなマルチレベル分析
 - 階層的線形モデリング (HLM)
 - マルチレベル構造方程式モデリング (ML-SEM)

マルチレベル分析のイメージ

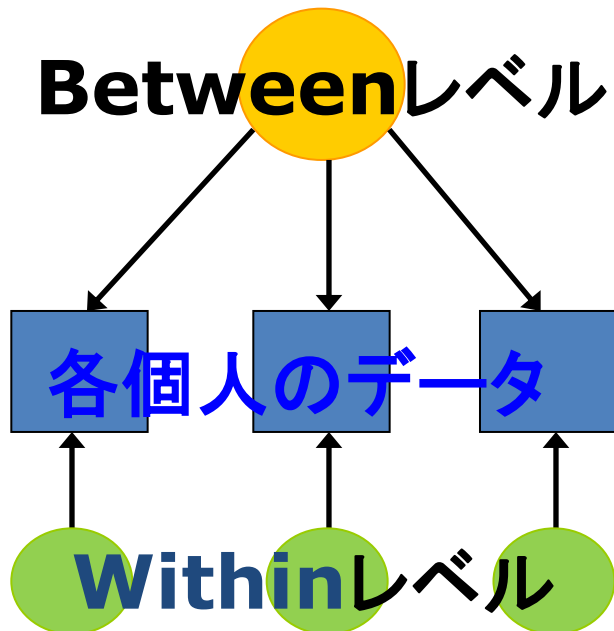
個々人のデータ(情報)

級内相関係数



マルチレベルモデルのイメージ

- モデルの仮定は因子分析と同じ
 - 因子分析: 行動傾向から、個人の態度を推定
 - MLモデル: 個人の得点から、集団の得点を推定



集団間変動

集団で共有された分散

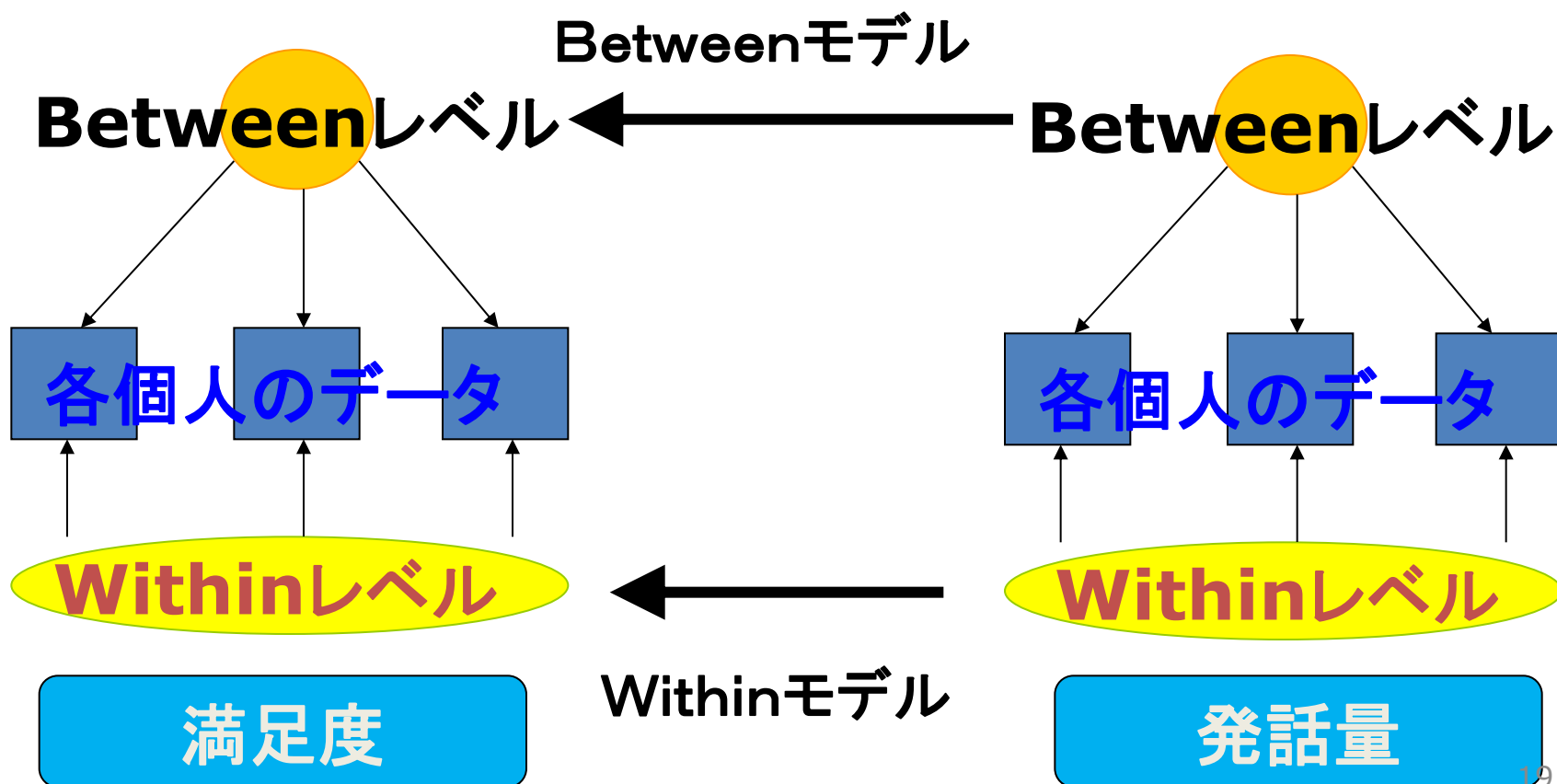
BetweenとWithinの分散をわけて、
それぞれの相関関係を推定する

集団内変動

個人独自の分散

マルチレベルモデルのイメージ

- 個人のデータから集団の効果を推定する



WithinレベルとBetweenレベル

- 各レベルごとにモデルを解釈
 - Withinレベル・・・個人間プロセスを表す
 - おしゃべりな人は、会話によく満足している
 - Betweenレベル・・・集団間プロセスを現す
 - よく話す集団は、みんなが満足している
- 各レベルの比較
 - 扱おうとしている現象は、個人間のプロセスなのか、集団全体のプロセスなのか？

マルチレベル相関係数

- Kenney & Lavoie (1985)が提唱
 - 普通の相関係数では混在してしまう, 集団間変動と個人間変動を分離して相関を推定できる
- Betweenレベルの相関係数
 - 集団内で類似している変動同士の相関
 - いわば, 集団を因子とした場合の因子間相関
- Withinレベルの相関係数
 - 集団の効果を統制した, 個人間変動の相関
 - いわば, 集団因子を取り除いた, 誤差間の相関

マルチレベル相関係数

- HADで分析
 - 満足度と発話量・・・Withinレベル < Betweenレベル
 - 集団成績と満足度も高い相関
 - 下三角行列がBetween, 上三角行列がWithin

マルチレベル相関分析		満足度	発話量	スキル	集団成績	条件
満足度	.358 **	.227 **	.290 **	.000	.000	
発話量	.466 *	.316 **	.113	.000	.000	
スキル	.000	.000	-.020	.000	.000	
集団成績	.498 **	.133	.000	1.00 **	.000	
条件	.315 *	.684 **	.000	-.097	1.00 **	
** $p < .01$, * $p < .05$, + $p < .10$						

相関分析 ※素点による相関係数							
		満足度	発話量	スキル	集団成績	条件	
	満足度	1.000					
	発話量	.307 **	1.000				
	スキル	.209 **	.141 *	1.000			
	集団成績	.297 **	.075	.053	1.000		
	条件	.188 **	.384 **	.048	-.097 +	1.000	
		** $p < .01$, * $p < .05$, + $p < .10$					

マルチレベル相関分析

		満足度	発話量	スキル	集団成績	条件	
	満足度	.358 **	.227 **	.290 **	.000	.000	
	発話量	.466 *	.316 **	.113	.000	.000	
	スキル	.000	.000	-.020	.000	.000	
	集団成績	.498 **	.133	.000	1.00 **	.000	
	条件	.315 *	.684 **	.000	-.097	1.00 **	
		** $p < .01$, * $p < .05$, + $p < .10$					

相関分析 ※集団平均値による相関係数

	満足度	発話量	スキル	集団成績	条件
満足度	1.000				
発話量	.371 **	1.000			
スキル	.118	.177 +	1.000		
集団成績	.394 **	.101	.095	1.000	
条件	.249 *	.522 **	.079	-.097	1.000

** $p < .01$, * $p < .05$, + $p < .10$

マルチレベル相関分析

	満足度	発話量	スキル	集団成績	条件
満足度	.358 **	.227 **	.290 **	.000	.000
発話量	.466 *	.316 **	.113	.000	.000
スキル	.000	.000	-.020	.000	.000
集団成績	.498 **	.133	.000	1.00 **	.000
条件	.315 *	.684 **	.000	-.097	1.00 **

** $p < .01$, * $p < .05$, + $p < .10$

マルチレベル相関からわかること

- 集団間プロセスと個人間プロセスの分離
 - 解釈の単位が，個人か集団かで異なる
 - 課題への満足度は，Withinレベルでは弱い相関，Betweenレベルでは中程度の相関
 - おしゃべりな人のほうが，満足する傾向がある
 - それよりも，みんなの発言量が多い集団みんなが満足する傾向のほうが強い
- 因果関係はわからない
 - →マルチレベルSEMへ

ここまでのまとめ

- 階層的データ
 - グループ内類似性＝級内相関が存在するデータ
 - データが独立にサンプリングされていないので、統計的検定に誤りが生じる
 - そのままでも、集団単位で平均しても、正しい推定値は得られない
- マルチレベル分析
 - 階層的データのグループ内類似性を評価
 - グループごとの類似性から、データを集団間変動と集団内変動に分離する
 - そうすることで、検定も推定値も正しい推定ができる

マルチレベル構造方程式モデル

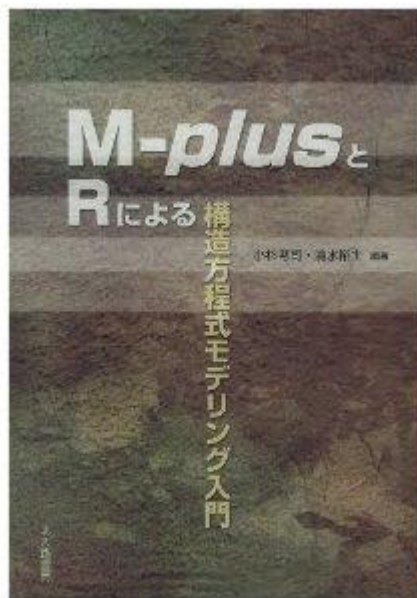
- 構造方程式モデルをマルチレベルに拡張
 - WithinとBetweenのモデルをそれぞれ推定
 - パスモデルが二つできる
 - 適合度は二つ合わせたものを参照できる
- 対応しているソフトウェアは少ない
 - Mplus・・・これがあれば大丈夫
 - マルチレベルアドインを入れる必要がある
 - Amos・・・とても面倒だけど、不可能ではない
 - 多母集団同時分析を使って、SEMの下位モデルとして表現
 - HAD・・・簡便法を近日実装予定
 - 来年の春あたり

Mplusがオススメ

- Mplusとは
 - Muthenらで作った, SEM用のソフトウェア
 - <http://www.statmodel.com/>
- 高度な統計解析が可能
 - カテゴリカル変数の分析や項目反応理論
 - マルチレベルモデル
 - 潜在混合分布モデル
 - またそれらを組み合わせた複雑なモデルも構築可能
- 無料ソフトを望むなら・・・
 - HADへの搭載をお待ちください

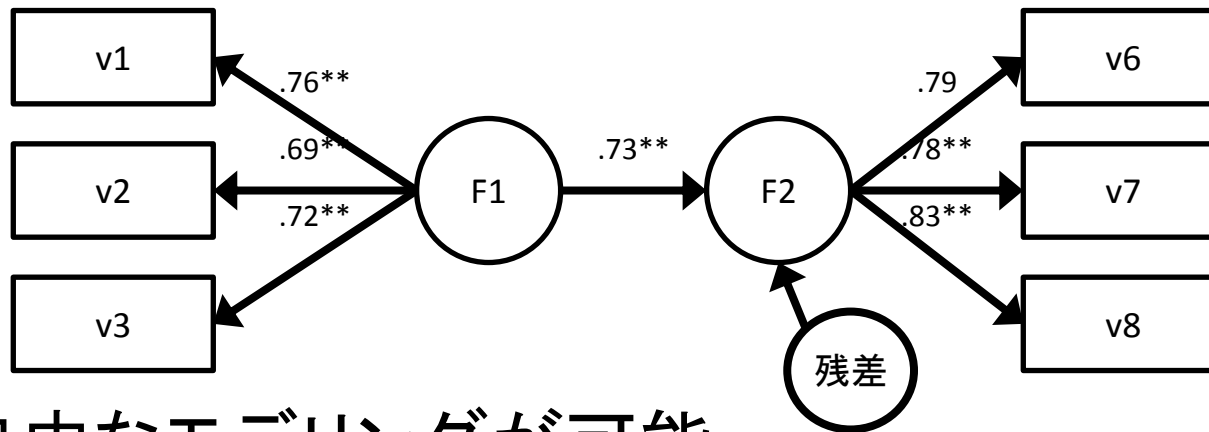
【宣伝】Mplusの日本語解説本登場【販促】

- 小杉考司・清水裕士 編著 (2014)
 - 『MplusとRによる構造方程式モデリング』
 - 北大路書房



SEMのおさらい

- 構造方程式モデル(SEM)とは
 - 潜在変数を含んだ因果関係を推定するモデル
 - 潜在変数を含まない場合は、パス解析と呼ぶこともある



- 自由なモデリングが可能
- 適合度指標を利用して、再現性を評価

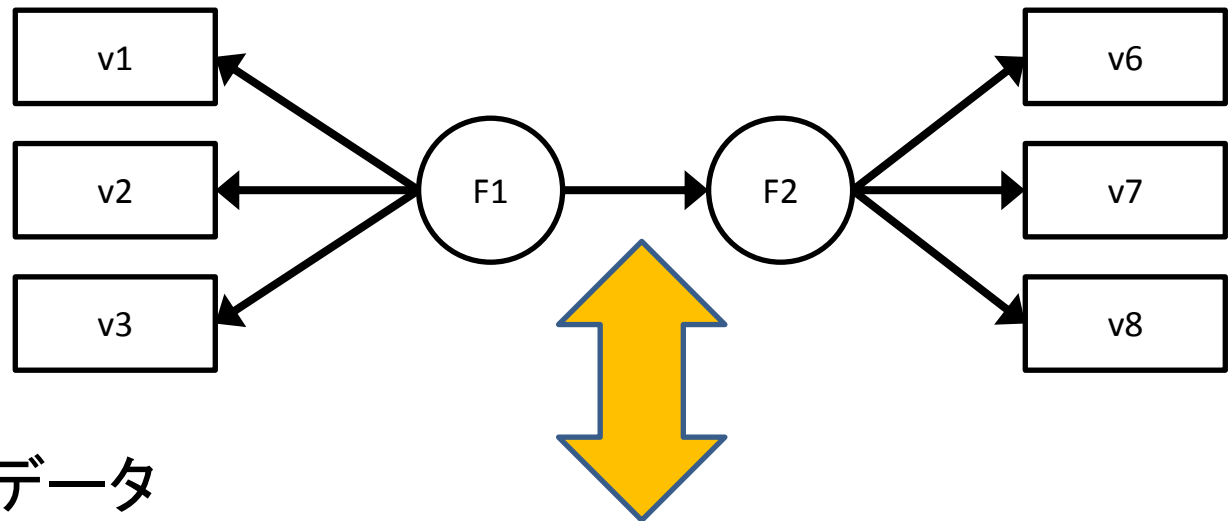
SEMがやっていること

- モデルを決める
 - パス, 共分散, 分散を指定
 - これらをパラメータと呼ぶ
- モデルを共分散構造に変換する
 - モデルを共分散行列の形に変換
 - SEMのデータは, 共分散行列
- 最尤法によってパラメータを推定
 - 共分散行列と共分散構造の距離が小さくなるように推定
 - 最尤法の場合, 距離のことを尤度と呼ぶ。

SEMがやっていること

- モデルと共分散行列の距離を小さくする

モデル



共分散行列 = データ

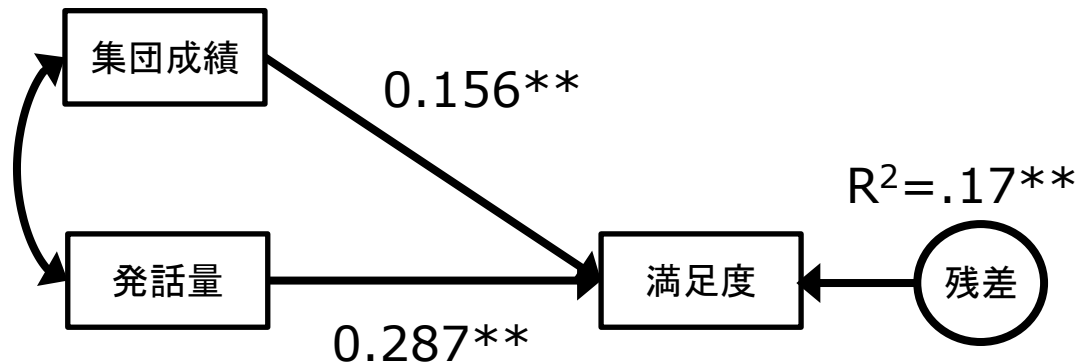
	v1	v2	v3	v6	v7	v8
v1	0.619874	0.320528	0.343342	0.270276	0.316181	0.345377
v2	0.320528	0.646206	0.335553	0.259271	0.234372	0.326181
v3	0.343342	0.335553	0.657663	0.260176	0.275377	0.306281
v6	0.270276	0.259271	0.260176	0.706508	0.450452	0.497035
v7	0.316181	0.234372	0.275377	0.450452	0.71196	0.476985
v8	0.345377	0.326181	0.306281	0.497035	0.476985	0.793869

適合度：データとモデルの距離

- SRMR・・・標準残差行列の二乗和の平方根
 - 最小二乗基準に基づく距離概念
- χ^2 乗値・・・尤度に基づいた距離
 - 検定もできるが、あまり精度はよくない
- CFI・・・一番悪いモデルに比べての相対的な適合度
 - 真のモデルを想定した場合の適合概念
- RMSEA・・・1自由度に対する適合度
 - モデルの規模に依存しない適合度
- 情報量規準
 - モデルの儉約性を考慮に入れた、比較のための適合度

SEMで推定

- 満足度を発言量と集団成績で予測



- 飽和モデルなので適合度は最大

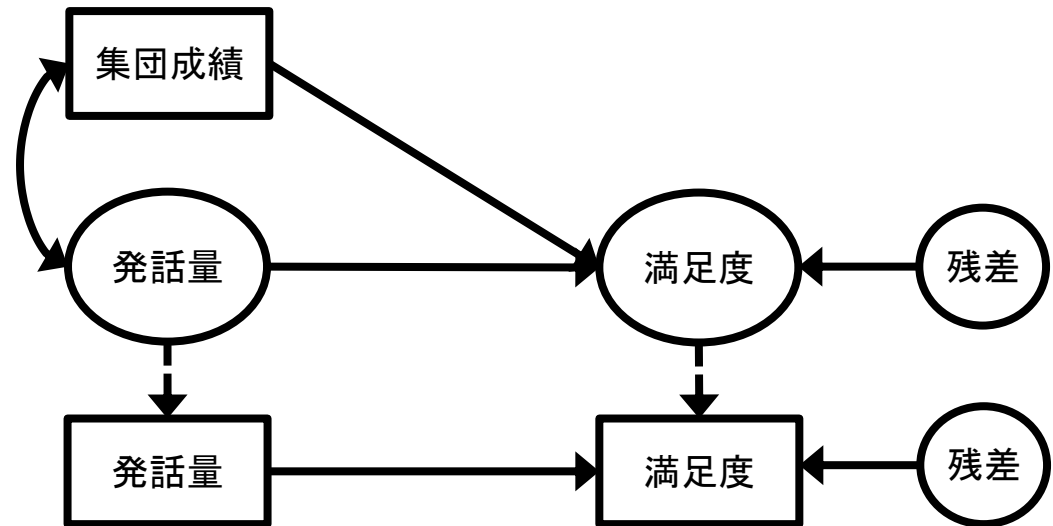
マルチレベルSEM

- SEMをマルチレベルに応用
 - 集団間変動のモデル = Betweenモデル
 - 集団内変動のモデル = Withinモデル
- モデルが二つ = データも二つ
 - Betweenの共分散行列とWithinの共分散行列
 - 二つのモデルを同時に最尤法で推定する
 - 実際には、データが二つに分離されているのではないが、イメージとしてはそんな感じ。

マルチレベルSEMのモデリング

- Between (集団)レベルを潜在変数として推定

– Betweenレベル

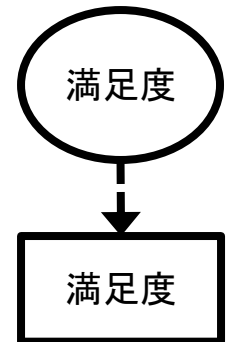


– Withinレベル

- ソフトウェア上では、モデルは別々に指定
 - 集団レベルが潜在変数になることは意識しなくてよい
 - 集団成績が顕在変数なのはもともと集団レベルの変数だから

ML-SEMのモデリングの意味

- 集団平均の集団間変動を推定
 - 例えば・・・課題満足度の得点
 - = 全体平均 + 集団の特徴 + 個人の特徴
 - 全体平均・・・定数
 - 集団の特徴・・・集団の数だけある(集団間変動)
 - 個人の特徴・・・個人の数だけある(個人間変動)
- Betweenレベル
 - 集団の特徴を取り出して, その集団間変動を予測する
- Withinレベル
 - 個人の特徴を残して, その個人間変動を予測する

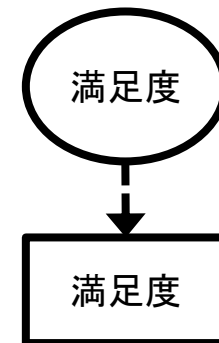


級内相関係数の確認

- 級内相関＝グループ内類似性
 - 級内相関が0より小さい場合，変数の得点に集団の特徴は含まれていない
 - 変数の得点 = 全体平均 + 個人の特徴
 - 級内相関が1の場合，個人の特徴は含まれない
 - 変数の得点 = 全体平均 + 集団の特徴
 - 級内相関が0～1の間の場合，両方の効果が含まれている

級内相関係数の推定

- Nullモデル・・・説明変数がないモデル
 - 純粋な目的変数の集団間変動を推定
 - それ以外にも, ソフトウェアで直接計算も可能
 - HADなど
- Mplusによる推定
 - 全体平均・・・3.433
 - 集団間変動・・・0.349**
 - 個人間変動・・・0.637**
 - 級内相関係数 = $0.349 / (0.349 + 0.637) = 0.354$ **



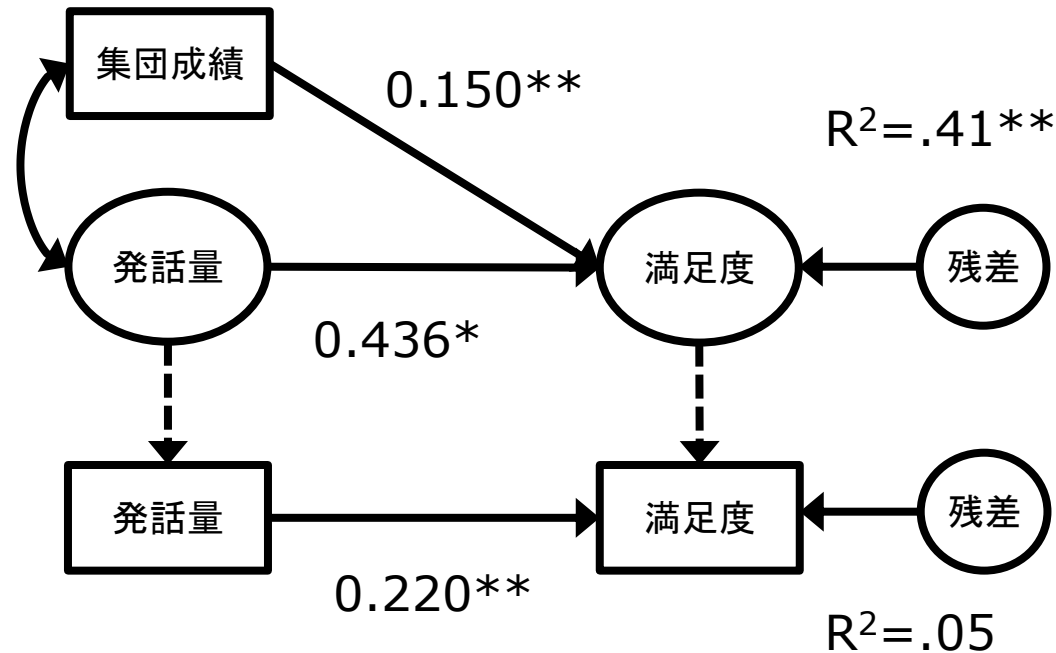
級内相関が極端な場合の対処

- 級内相関が0以下, あるいはかなり小さい
 - Withinレベルのみで用いる
 - その場合, 得点を集団平均で引いておくとよい
 - 集団平均中心化
- 級内相関が1, あるいはかなり1に近い
 - Betweenレベルのみで用いる
 - もし, 級内相関が完全に1ではないなら, 集団平均値を用いる

回帰モデルを推定 (Mplus)

- Betweenレベル

切片の集団間変動: 0.349 → 0.205



- Withinレベル

切片の個人間変動: 0.637 → 0.604

Withinモデルの説明率よりも, Betweenモデルのほうが高い!

結果の解釈例

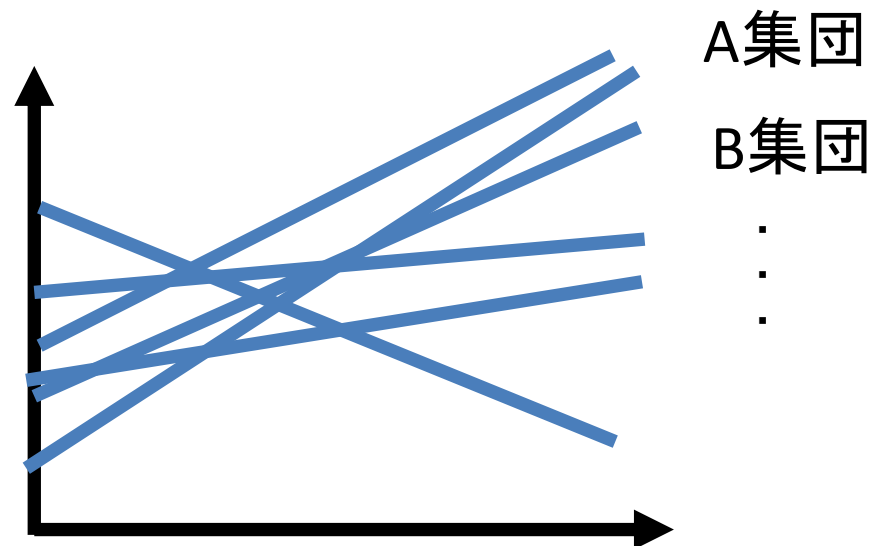
- 課題の満足度は、Betweenモデルで説明
 - よくしゃべった人が、課題に満足するというより、よく話している集団は、みんな満足が高くなる
 - 自分がただ話す効果よりも、自分の発話に対してレスポンスがある、あるいは、会話が継続的に行われていることのほうが重要である可能性
 - 集団成績も、満足度に影響
 - 課題成績が良かった集団は、みんな満足
- Withinレベルも無視はできない
 - 発話することに効果も若干ある
 - スッキリする、イニシアチブをとる、などなど

ここまでのまとめ

- マルチレベルSEM
 - 集団平均の集団間変動を潜在変数として推定
 - 個人の得点 = 平均 + 集団効果 + 個人効果
 - 集団効果と個人効果をそれぞれモデリング
 - Betweenレベル
 - Withinレベル
- 級内相関係数の推定
 - 級内相関が極端な場合は、対処が必要
 - Nullモデルを推定
 - 目的変数だけのモデル
 - 最初のモデルからそれぞれの変動がどれほど説明されているかを確認する

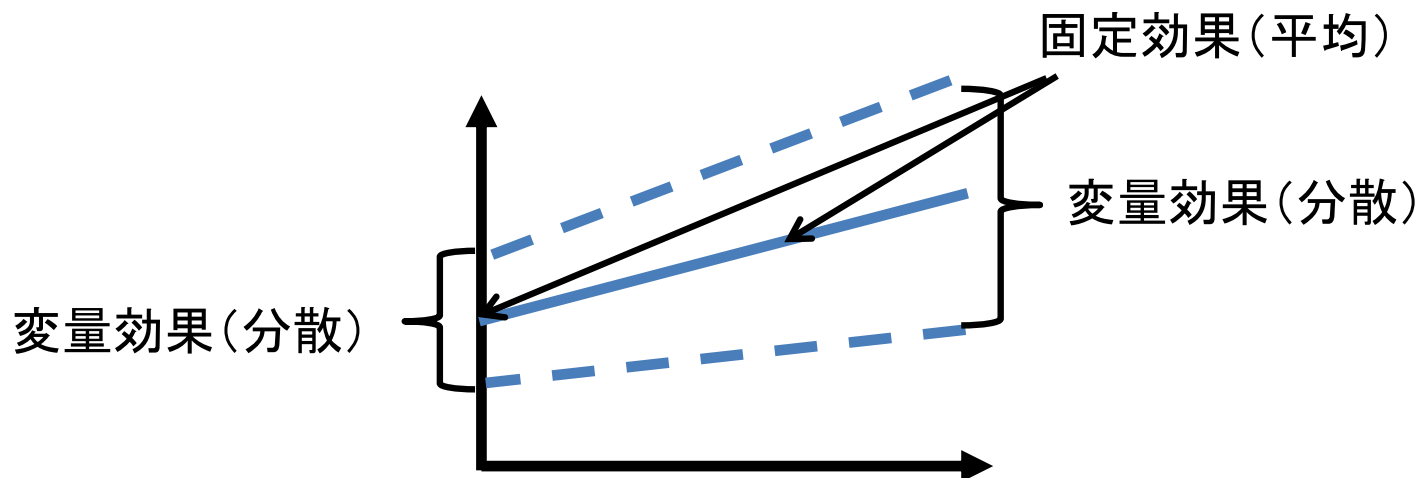
Withinのパス係数の集団間変動

- パス係数をグループごとに推定すると・・・
- 100グループの場合、パス係数は100個
 - 中にはパス係数が大きいグループもあれば、小さいグループもあるかもしれない

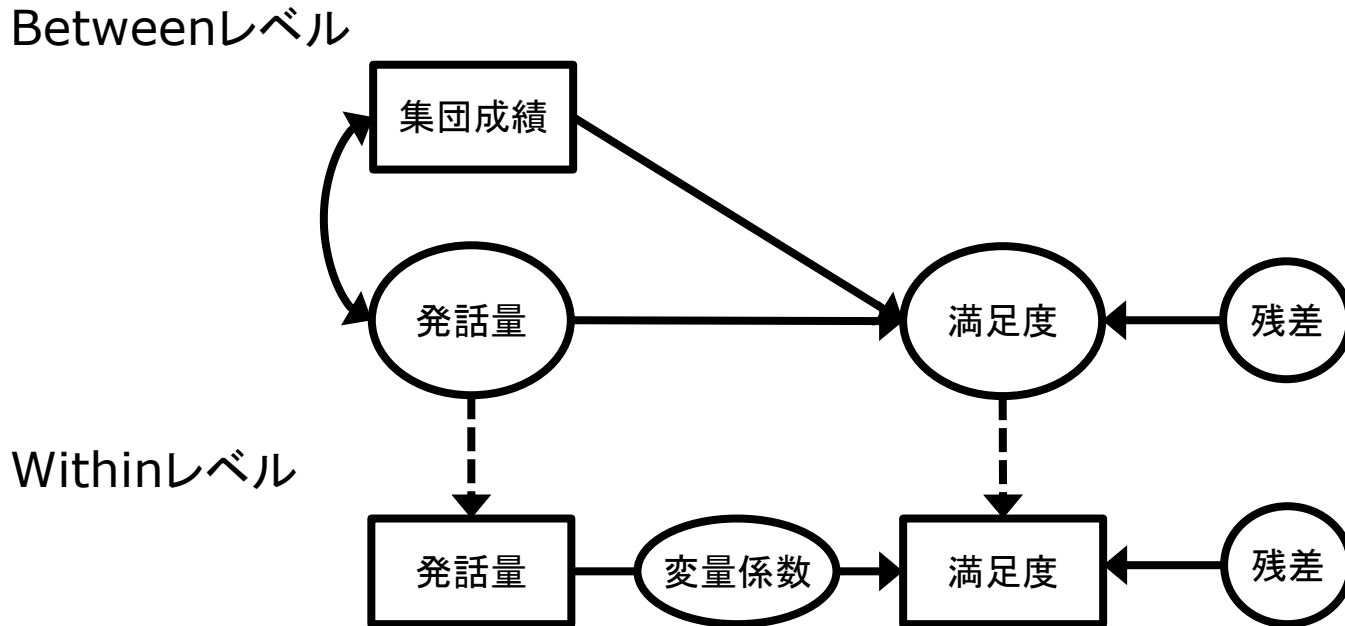


Withinのパス係数の集団間変動

- パス係数の集団間変動を推定する
 - パス係数の平均と分散が推定できる
 - パス係数が確率的に変動する(正規分布)と考える
 - 集団全部の平均値・・・固定効果(ただのパス係数)
 - 集団で変化する値・・・変数効果(係数の集団間変動)



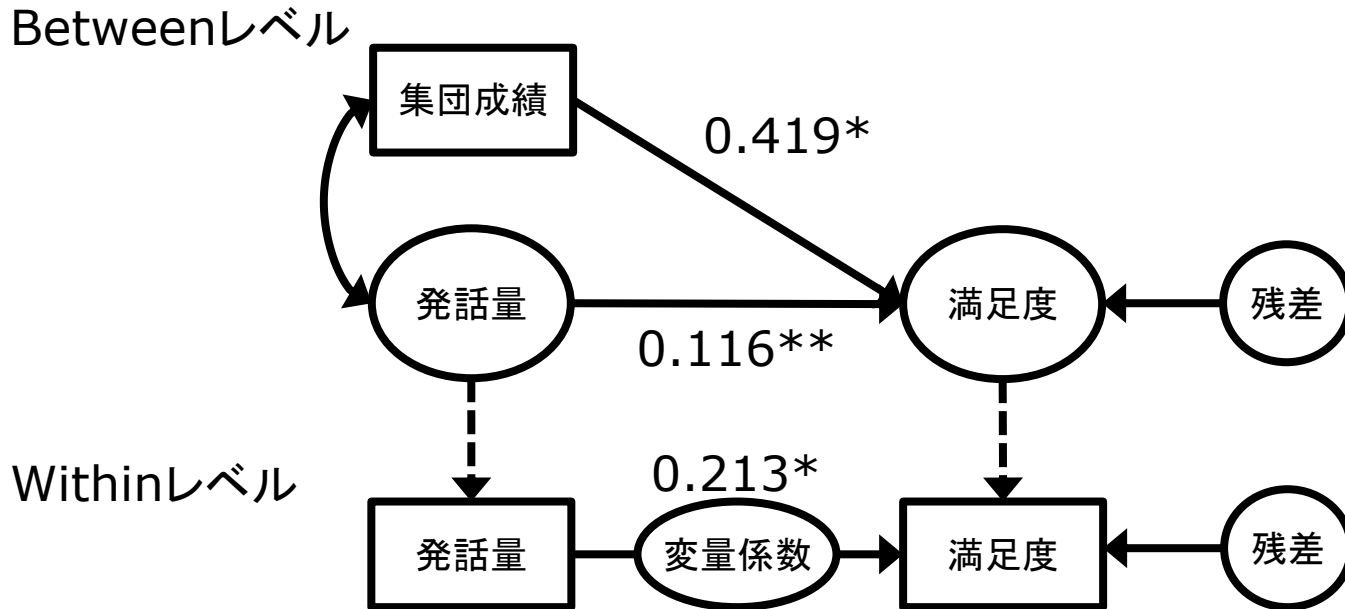
パス係数の集団間変動をモデル化



変量効果を仮定したパス係数 = 変量係数
パス係数の集団間変動を分散成分で表現

発言量と満足度のWithinレベルの関連が、集団ごとによって異なることを仮定

Mplusで推定



変量効果の分散成分(集団間変動) = 0.219^{**}

変量効果の分散成分が有意・・・集団間で係数に違いがある！

パス係数の変動を予測する

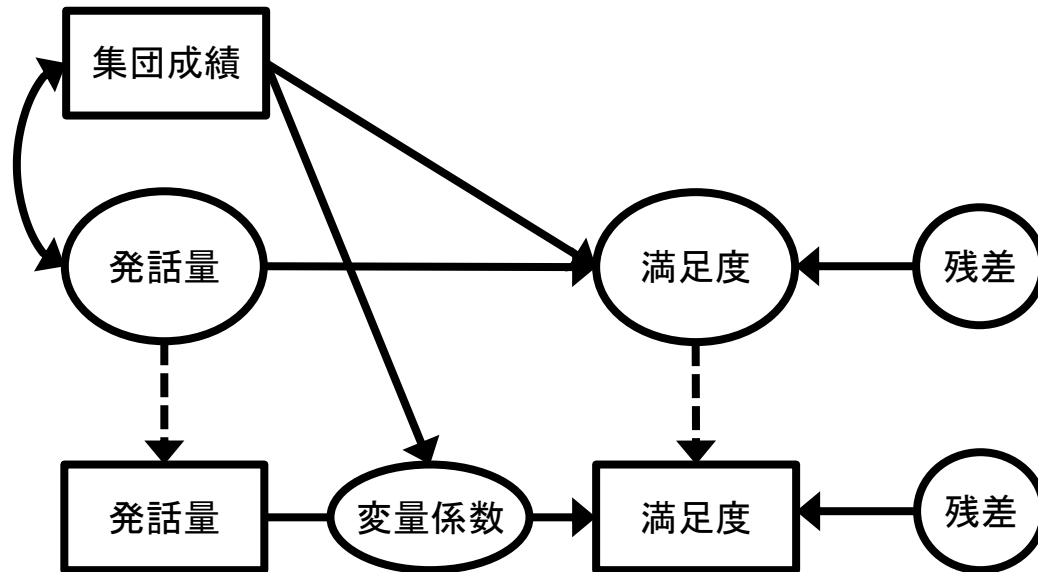
- パス係数の集団間変動がある場合
 - 集団の性質によって、Withinレベルのパス係数が異なる可能性が示唆される
 - もし、その集団の性質が、Betweenレベルの変数によって説明できるのであれば……
 - パス係数の集団間変動を、Betweenレベルの変数で予測するモデルを構築できる

パス係数の集団間変動を予測する レベル間交互作用

- 回帰式を次のように表現する
 - 満足度 = $\beta_0 + \beta_1 \times \text{発言量}$
 - $\beta_1 = \gamma_{10} + \gamma_{11} \times \text{集団成績}$
 - この式を上での β_1 に代入すると...
 - 満足度 = $(\gamma_{10} + \gamma_{11} \times \text{集団成績}) \times \text{発言量}$
 - この式を展開すると...
 - 満足度 = $\gamma_{10} \times \text{発言量} + \gamma_{11} \times \text{集団成績} \times \text{発言量}$
 - つまり、パス係数の集団間変動を予測する式は、交互作用項と同じように、係数同士の積の項であることがわかる
 - これをレベル間交互作用、と呼ぶことがある

パス係数の集団間変動を説明する レベル間交互作用

Betweenレベル



Withinレベル

仮説(例):

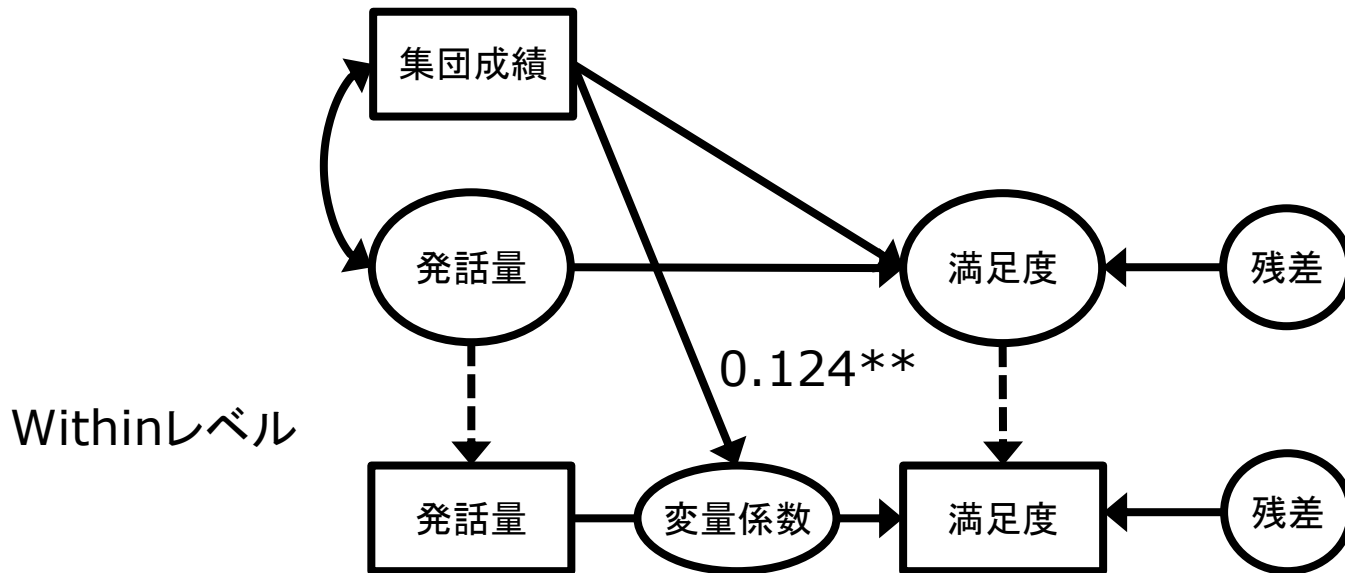
集団成績が良い集団は、発話量が多いほど満足するが、
逆に集団成績が低い集団は、関連がない

レベル間交互作用を見る場合の注意

- 得点の中心化を行う
 - Betweenレベルの変数の平均値を0にする
 - 全体平均を得点から引けばよい
- なぜ中心化を行うのか？
 - 解釈のしやすさのため
 - Withinレベルの回帰係数の固定効果は、切片なので、説明変数が0のときの値になる → 中心化すると解釈しやすい
 - 多重共線性を小さくするため
 - 中心化しないと、主効果と交互作用効果の相関が高くなり、標準誤差が不当に大きくなる

Mplusによる推定

Betweenレベル



変量効果の分散成分(集団間変動)=0.219→0.158(約28%説明)

変量係数に有意な効果:

集団成績は発話量と満足度の関連を調整していることがわかる

結果の解釈例

- Withinレベルの効果は集団によって違う
 - (当たり前だが)発話内容が違っている可能性
 - ただ単に量だけに注目することの問題
- 集団成績によって、発話の効果は調整
 - 成績が良かった集団
 - ポジティブな内容の発言が多いかもしれない
 - 成績が悪かった集団
 - グチをこぼしたり、自分の発言が報われなかったことによる不満感かもしれない

推定についての注意点1

- WithinレベルとBetweenレベルの推定値
 - 片方が変われば, もう片方も推定値が変わる
 - 完全に独立なわけではなく, 双方に影響している
- 標準化係数の解釈
 - 得点を標準化したものとは一致しない
 - 通常のSEMとは大きく異なるところ
 - 各レベルで標準化されている
 - 級内相関が小さい場合は, Betweenレベルの標準化係数は不当に大きくなることもある
- 変量係数を仮定すると標準化できない
 - 目的変数の分散が厳密に定義できないから
 - 分散説明率も, 負になることがあり, うまく定義できない

推定についての注意点2

- 不適解が多い
 - 分散成分が負になってしまい、推定がうまくいかないことも多々ある
- Betweenレベルのサンプルサイズ
 - 20ぐらいはないと、不安定な結果
 - しかし、実は集団の数よりも、集団内人数のほうが推定の安定性には影響する
 - ペアデータとかだと、ペア数が多くても推定が不安定(標準誤差が大きくなる)になることがある

ML-SEMをベイズ推定で行うメリット

- 最尤法の欠点である「不適解」を回避
 - レベル2の推定値は不安定なので最尤法はきつい
 - ベイズは非負の制約を簡単にかけられる
- 推定値の正規性の仮定を必要としない
 - 一般に、分散成分は正規分布しない
 - ベイズのほうがデータにあった推定が可能
- 小さいサンプルサイズでもそれなりに推定できる
 - 最尤法は漸近理論に基づく→大標本が必要
 - 集団は数多くサンプリングしにくい →ベイズに利点
 - ただし、適切な事前分布を仮定する必要がある
- Mplusがあれば、ベイズ推定も簡単

結果の書き方

- 基本は普通のSEMと同じ
 - 推定値, 標準誤差(信頼区間もあるとよい)
 - ただ, 分散成分の信頼区間はブートストラップのほうがよい
 - 適合度指標(χ^2 乗値, CFI, RMSEA, SRMRなど)
 - 変量係数がある場合, 情報量規準のみが推定される
 - パス図もあるとよい
 - パス図はレベルごとに分けて書くほうがわかりやすい
- マルチレベル分析では分散成分が重要
 - 集団間変動を表す分散成分は必ず表記
 - 変量切片, 変量係数ともに

レベル間の関係と解釈

- WithinレベルとBetweenレベルの関係
 - 個人が集団に影響することはないの？
 - マルチレベルモデルでは、想定しない
 - なぜか
 - 厳密には、個人レベル \neq Withinレベル
 - Withinレベルは、Betweenレベルを統制した残差の効果
 - 個人全体の情報は、測定された変数そのもの
 - 個人が回答した情報には、集団全体で共通する変動と、それ以外の変動に分離できる・・・それ以外=Within

レベル間の関係と解釈

- 概念的な話と分析的な話の違い
 - 個人・集団というレベル分けは概念的な話
 - その分け方は研究者の視点に依存する
 - 発話を個人のものとするか, 複数人の相互作用ととらえるか
 - Within・Betweenというレベル分けは分析的な話
 - 集団内変動と集団間変動を区別する
 - 発話量の集団平均の変動と, 集団平均からの偏差の変動

例：文化比較の研究

- 社会関係資本と個人の主観的幸福感
 - 国ごとの主観的幸福感の違いを社会関係資本で説明する
 - Betweenレベルで効果があったとき
 - 国の幸福感を社会関係資本で予測？
 - 国の主観的幸福感って何？
 - 国民それぞれの主観的幸福感が、底上げされていると考えたほうが自然
 - あらゆる概念を集団レベルの概念として考えられるわけではない
 - それは研究者の立場による

例：文化比較の研究

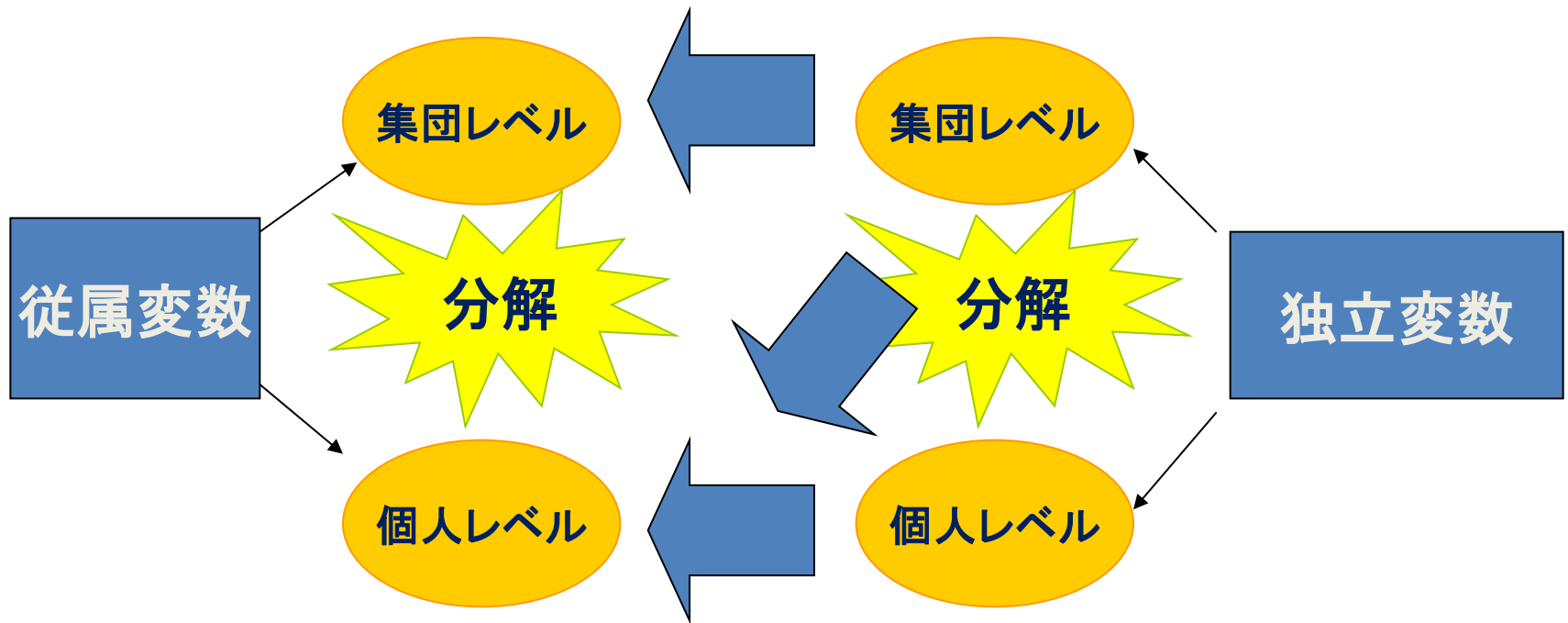
- 社会関係資本と個人の主観的幸福感
 - 国ごとの主観的幸福感の違いを社会関係資本で説明する
 - Withinレベルで効果があったとき
 - 個人の社会的関係資本が主観的幸福感を予測？
 - 個人の社会的関係資本って何？
 - 概念的に定義できるのか、あるいはただの個人の認知バイアスなのかを、しっかり考える必要がある
 - 集団レベルと同様、Withinレベルの分散が、そのまま個人レベルの概念として解釈わけではない
 - 研究者の立場や理論次第

他の手法との違い

- HLMと比べて
 - モデリングの柔軟さでは、マルチレベルSEMが上
 - HLMはML-SEMの下位モデルに位置付けられる
 - しかし、目的変数が一つだけなら、細やかな設定ができる点で、HLMのほうが便利なこともある
 - 分散成分の推定は、HLMのほうがバイアスが小さい
 - HLMは対応しているソフトウェアが豊富
 - ユーザーの手が届きやすいのはHLM
- 多母集団同時分析と比べて
 - 使いやすさは集団の数に依存する
 - 集団数が極端に少ない(5以下?)なら、多母集団
 - 集団数が20程度あるなら、マルチレベルモデル
 - その間・・・一番厄介

ML-SEMとHLMのイメージ

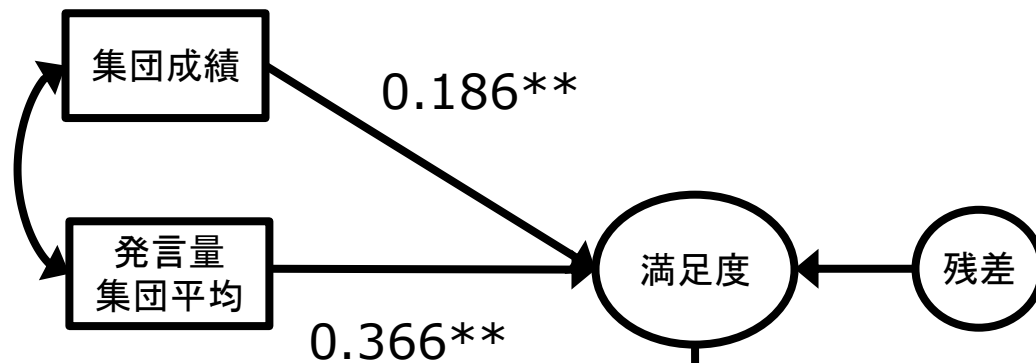
ML-SEM



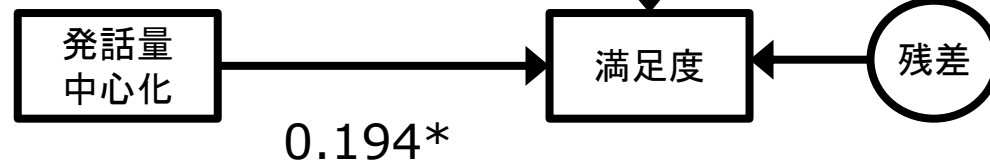
集団レベルのモデルは、
ML-SEMのほうがより正確に分析できる

HLMでの推定結果

Betweenレベル



Withinレベル



まとめ

- 階層的なデータに対する解決法
 - 従来法では, 検定結果と推定の解釈を誤る
 - マルチレベルモデルを用いれば, その両方を解決することができる
- マルチレベルSEM
 - WithinとBetweenレベルに分けてモデリング
 - レベルごとに解釈を行う
 - 解釈の仕方は難しいが, まずはレッツトライ!

参考になる文献

- 狩野裕・三浦麻子 2002
– グラフィカル多変量解析増補版 現代数学社
- 豊田秀樹 2000
– 共分散構造分析 応用編 朝倉書店
- 小杉考司・清水裕士 2014
– MplusとRによる構造方程式モデリング 北大路書房