

日本心理学会第87回大会 公募シンポジウム

## 心理学研究における構造方程式モデリング (SEM) の応用のあり方：双因子モデル、縦断データ解析、有限混合分布

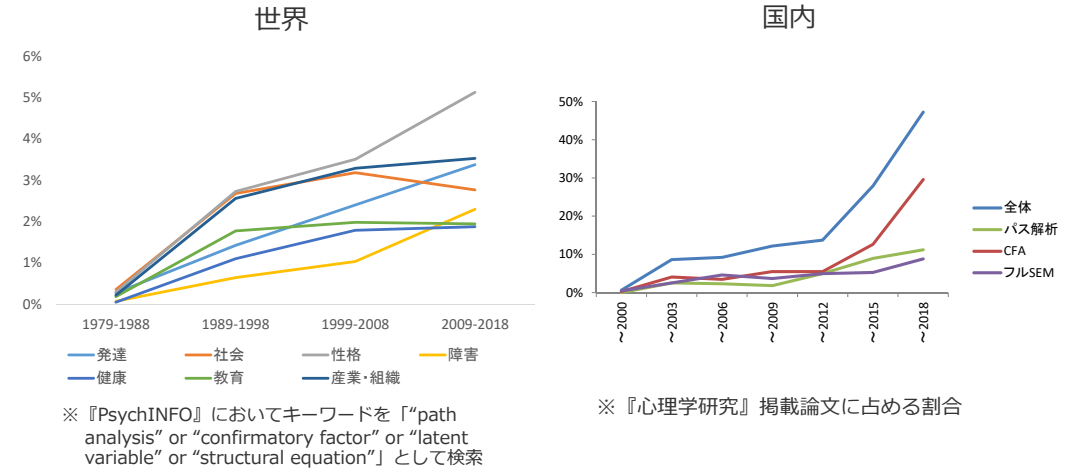
企画・司会・話題提供：伊藤大幸（お茶の水女子大学）

話題提供：谷伊織（愛知学院大学）

話題提供：村上隆（中京大学）

指定討論：山田剛史（横浜市立大学）

## SEMの利用動向



## 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カバー
  - ✓ 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）
  - ✓ 成長曲線モデル
  - ✓ マルチレベルモデル
  - ✓ 混合分布モデル 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との付き合い方

- ◆ 最も誤用されやすい手法でもある
  - Big numbers, low quality (Kline, 2023)
  - 権威ある学術誌に掲載された論文にも、モデル設定や解釈の重大な誤り (Zhang et al., 2021; 伊藤, 2019)

ユーザーである心理学研究者自身が、「マニュアル依存」の盲目的な解析に陥ることなく、SEMの研究応用のあり方を主体的に考え、議論していく必要がある

## 企画趣旨

- ◆ 2014年度の大会から継続的にSEMに関する企画を開催
- ◆ 初学者向けのテキストも刊行
- ◆ テキストの発展篇を執筆する中で見えてきた論点を指定討論者や聴衆と共有し、議論を深めたい



## 本日の内容

- ◆ 話題提供
  - 量的変数から得られるカテゴリカルな潜在変数：村上隆（中京大学）
  - 縦断データ解析におけるSEMの応用：伊藤大幸（お茶の水女子大学）
- ◆ 指定討論
  - 山田剛史（横浜市立大学）

量的変数から得られるカテゴリカルな潜在変数：

臨床診断におけるdimensional systemと  
categorical systemの「対立」を題材に

村上 隆  
(中京大学・文化科学研究所)

## 本報告の概要

- 有限混合分布とは
- 臨床診断への応用可能性
- 有限混合分布の妥当性： シミュレーションによる検討
- 最近の発展の一例： 分布仮説とサンプルの比率
- 結論

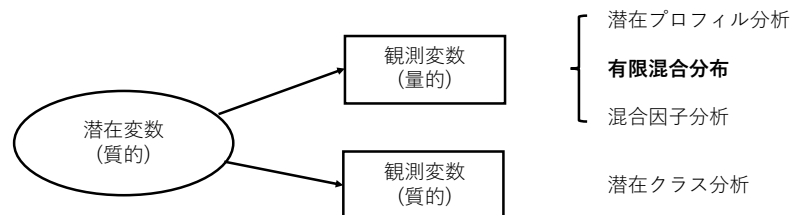
## 有限混合分布とは

## 有限混合分布

- 量的・質的データから量的な潜在変数を導出し、その間の（因果）関係を明らかにしようとする試みは、構造方程式モデルとして、心理学研究では定着している
- 他方、ここでとり上げるのは、量的変数からカテゴリカルな潜在変数を導出しようとする、一般に Mixture modeling と呼ばれている方法についてである。
- この方法について、特に臨床診断の分野での適用例を題材に、その意義と問題点を探ろうとするのが本提案の目的である。

## Mixture modeling と呼ばれる一連の方法群

- 歴史的源泉はいくつかある (finite mixture distributions, latent structure analysis, latent class analysis, latent profile analysis, etc.) が、現時点では、**潜在変数としての独立変数が、(量的でなく) カテゴリカルに与えられる構造方程式モデリング**とみなすことで、統一的なアルゴリズムが構築されている (Mplus)。



## なぜ Categorical な潜在変数なのか？

- 個人差を指示する多くの概念には、明確なカテゴリーの境界が定義できるものはめったにない。
- そうした概念の測定を目指す尺度の得点 (分布) にも、多くの場合、そうした境界スコア (cut point) を明示するような特徴を備えたものは、ほとんどない。
- しかしながら、「人間は分類が好きである (Frances, 2013)」
- さらに、測定 (診断) の対象となった個人の処遇を連続量にしたがって、いわばきめ細かく連続的に定めて実行できることはほとんどない。

## ROCカーブ：検査スコアの感度と特異度

- 個人の検査の得点と、同じ個人の疾病の有無についての (確実な) 診断結果が存在するとしたとき、検査得点ごとに、実際には疾病がなかった比率 (特異度) を横軸に、疾病があった比率 (感度) を縦軸にプロットしたもの (あるいは点をつないで描いた曲線) をROCカーブという。
- 得点上に適切な cut point を定めることにより、最適な (と考えられる) 感度と特異度をもつ (検査得点だけにもとづく) 診断基準を定めることができる (後に例を示す)。

## 検査得点の分布に診断の手掛かりはないか？

- しかし、必ずしも確実な診断基準が確立していない、あるいはそれに著しいコストのかかる場合、検査スコアのみから、syndrome の有無を判定できる方法はないか？
- **検査スコアの分布自体に、カテゴリカルな診断の手掛かりがあるのではないか？**
- 二峰性の分布があれば、中間の谷のところに、比較的あいまいさのない cut point を決めることができる (かもしれない)。
- そうでない場合も、正規分布ではない偏った分布は、複数の正規分布の混合 (mixture) からなっているのかもしれない。
- だとすれば、度数分布から、混合する正規分布の割合とパラメータを推測できないか？ それが可能なら、個人のスコアにもとづいて、診断を行うことができるのではないか。

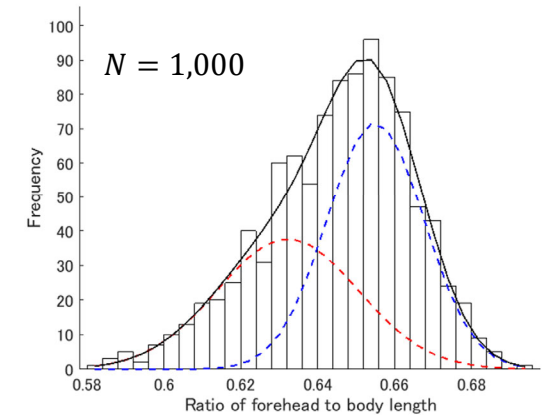
## 本報告における，有限混合分布

- 本提案では，基本的に1変量の度数分布の中に，複数の集団を見分けるという，19世紀の Carl Pearson に起源をもつ方法に限ってやや掘り下げて検討し，その多変量への展開については今回は触れないことにしたい。
- **単純なモデルに絞って議論しようという考えが，あまり良い知恵でないことが後程わかりますが・・・。**

## Carl Pearson (1894) のカニのデータ

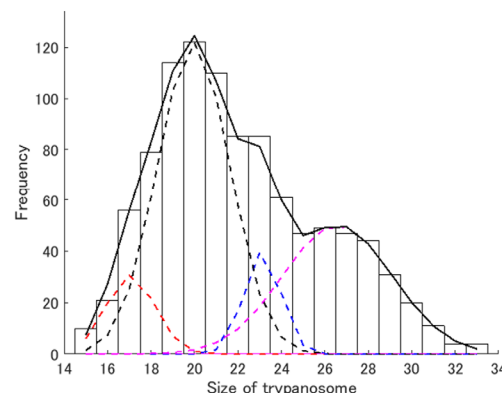
- 19世紀にすでに考えられていた有限混合分布。
- 右は，Mplus 8.10 で計算された最尤解だが，Pearson は moment 法によって，ほぼ同じ解に到達していた。

	Mean	SD	Proportion
Group1	0.632	0.0183	0.433
Group2	0.655	0.0126	0.567
Total	0.645	0.0191	



## 3つ以上の正規分布が必要な場合も

- こちらは，複数の感染症を媒介する trypanosome という原生動物の大きさの分布である (Pearson, 1914)。
- かなり複雑な形状を呈しているが，それでも4つの正規分布の合成でかなり正確に記述できる。
- 大きさにもとづいて，各個体がどの「種」に属するかを判定できる (とされる)。

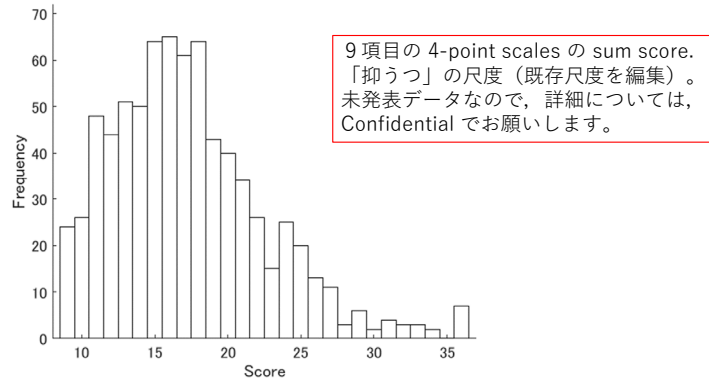


**ここで見出されたのは4つのカテゴリカルな値をもつ1つの潜在変量**

## 臨床診断への応用可能性

## ある心理臨床家への一言

- ・村上「実務家にとっては、正規分布しないスコアは使いにくいかもしれないですね」



## 心理臨床家からのコメント

臨床現場でスクリーニングやアセスメントで尺度を使用した場合、

1. この人は、うつ病に当てはまるのか？
2. この人の症状は、どのくらい重いのか？

という情報が欲しいと思います。

- 1は「カットオフポイントはどこになるのか？」になります。
- 2は「（特に高得点者の）個人の得点が分布のどの位置にいるか？」になります。

## 続けて同じ人からの質問

- ・これらの2つについて正しい情報を得るためには、スコアの正規分布は必要でしょうか？
- ・低得点側に分布の偏りがある尺度で、前述の1と2の情報を得ようとするのは不適切でしょうか？

## さらに続けて・・・

- ・（一部の人が該当する）不健康度を測る尺度を、不健康な人も健常者も含めて全体に実施すれば、当然分布は偏ります。
- ・本研究のように健常者のみであれば、なおさらです。
- ・前述の1と2の情報が正規分布しないスコアからも得られるならよいですが・・・
- ・いかがでしょうか？

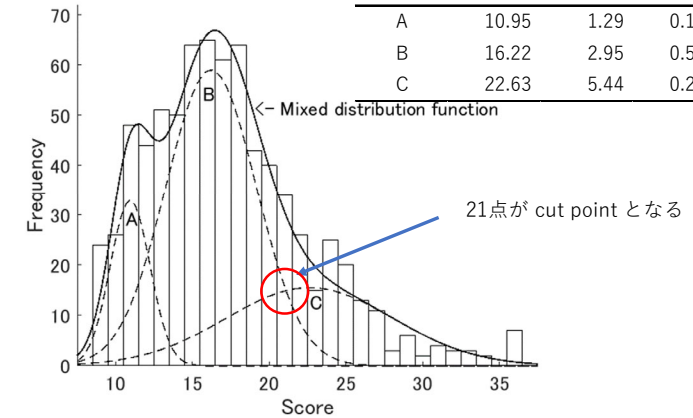
# 混合分布は活用できるか

- Cut point を有限混合分布によって、定義することはできるだろう。
- 有限混合分布で、健常者群と障害者群の分布が得られれば、それぞれは、正規分布となる。症状の重さは正規分布における相対的位置（標準得点）で示せる。
- それは、実務家の目的にかなう representation になるのだろうか。

- 3つのクラスが必要のようである。
- Empirical distribution への適合はますます良好か

表1 3つの正規分布の記述統計量と割合

	平均値	標準偏差	割合
A	10.95	1.29	0.141
B	16.22	2.95	0.578
C	22.63	5.44	0.281



# Bayes の定理に基づく割り当て

表2 素点ごとの各クラスへの帰属確率とクラスへの割り当て

得点	確率			割り当て	度数	得点	確率			割り当て	度数
	A	B	C				A	B	C		
9	0.746	0.207	0.048	A	24	22	0.000	0.358	0.642	C	26
10	0.772	0.195	0.032	A	26	23	0.000	0.212	0.788	C	15
11	0.703	0.263	0.034	A	48	24	0.000	0.107	0.893	C	25
12	0.501	0.450	0.049	A	44	25	0.000	0.047	0.953	C	20
13	0.207	0.721	0.072	B	51	26	0.000	0.018	0.982	C	13
14	0.040	0.873	0.087	B	50	27	0.000	0.006	0.994	C	11
15	0.004	0.899	0.097	B	64	28	0.000	0.002	0.998	C	3
16	0.000	0.888	0.112	B	65	29	0.000	0.001	0.999	C	6
17	0.000	0.862	0.138	B	61	29	0.000	0.001	0.999	C	2
18	0.000	0.820	0.180	B	64	30	0.000	0.000	1.000	C	4
19	0.000	0.752	0.248	B	43	31	0.000	0.000	1.000	C	3
20	0.000	0.652	0.348	B	40	32	0.000	0.000	1.000	C	3
21	0.000	0.516	0.484	B	34	33	0.000	0.000	1.000	C	2
						34	0.000	0.000	1.000	C	0
						36	0.000	0.000	1.000	C	7

(右の表に続く)

# 次のような判断ができる

- Class の分布が正規分布であることから、ある個体が属する Class が決まれば、その中での相対的位置は簡単に計算できる。
- また、score を偏差値に換算しても、（正規分布と偏差値の関係についての知識があれば）「正しく」解釈できる。

## 具体的には・・・

- たとえば、24点の人がいたとすると、表2から、Class C に属する確率が 0.893, Class B に属する確率が 0.107 である。
- 単純にわりあてれば、Class C つまり障害者群に属することになる。この場合、標準得点で、

$$z = (24-22.63)/5.44 = 0.252$$

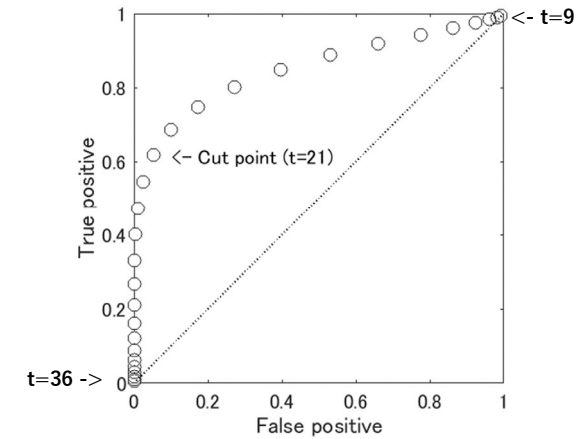
偏差値なら

$$Z = 10 \times 0.252 + 50 = 52.5$$

- これは、それほど極端な位置ではない。Class B に属するとしたら、 $z = 2.641$ 、偏差値では約 76 でこれは健常者としては上限に近い位置となる。

	平均値	標準偏差	割合
A	10.95	1.29	0.141
B	16.22	2.95	0.578
C	22.63	5.44	0.281

## ROC曲線：感度よりも特異度？



## いろいろと疑問は湧いてくる

- この分析は、28個の得点に対する度数分布だけに基づいている。
- つまり、分布の形状、特に歪度に全幅の信頼を置いている。
- 分布の偏りが大きいほど、障害群が増加することになるが、それでよいのか。
- Class C が28.1%も占めるのは常識的に考えて多すぎないか。
- Class C の分散は大きすぎないか（感度が低い理由）。換言すれば、そんな低得点のところまで、「隠れ鬱」がいるのか。

## 有限混合分布の妥当性

### シミュレーションによる検討



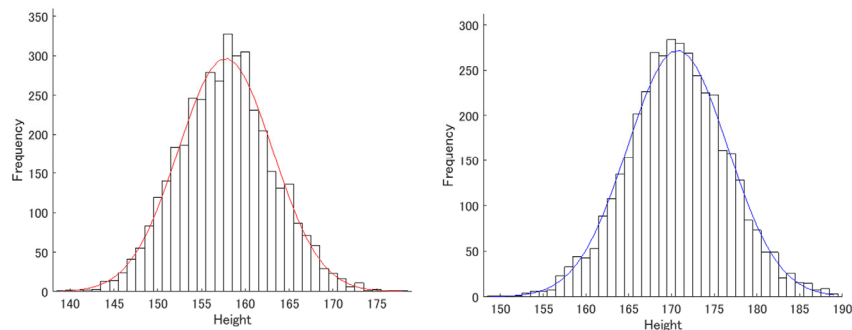
## 適用したデータ

- もっとも基本的と考えられる 1 変数, 2 クラスのケースに限定する。
- 文部科学省による学校保健調査 2019年版より, 17歳男女の身長と体重のデータ。身長は cm 刻み, 体重は kg 刻みの千分率で表示されているが, これを母集団とみなし, 累積百分率 (小数点以下 1 桁) に直した上で, 一様乱数を用いた疑似的サンプリングによって, それぞれから任意の大きさのデータを生成する。身長は, 性別ごとに見ると, ほぼ完全な正規分布であり, 体重の方はかなり正に歪んだ分布である。
- 身長の平均値  $\mu$  と標準偏差  $\sigma$  は, 女子で  $\mu = 157.8, \sigma = 5.3$ , 男子は  $\mu = 170.6, \sigma = 5.8$  である。

## 男女の身体計測を例に用いる意味

- 混合分布の研究の歴史は長く, 最尤法の適用についても, たとえば, EM algorithm の最初期の Example に含まれているところからみても, 多くの経験が蓄積されているものと思う。
- しかし, 実データにおいて, 潜在クラスが既知であるようなものは少なく, 実在する潜在クラスがどの程度正確に再現できるかという観点からの研究例は, あまり見られない。
- また, 混合分布モデリングには, 適合度指標が作りにくく, 情報量基準等による複数のモデルやクラス数間の相対的比較にとどまらざるを得ない。その点でも, 既知の潜在クラスの再現性という基準は, 混合分布モデリングの妥当性検討となる可能性があると思われる。

## 17歳の身長はほぼ正規分布に見える



女子 (左) と男子 (右)。それぞれ  $N = 4000$  のサンプル

## シミュレーション：標本の特異性

### • 標本の記述統計量

		標 本									
		1	2	3	4	5	6	7	8	9	10
平均値	女子	157.7	157.7	157.8	157.8	157.7	157.9	157.8	157.8	157.9	157.8
	男子	170.7	170.5	170.7	170.7	170.6	170.5	170.6	170.6	170.6	170.5
標準偏差	女子	5.38	5.30	5.26	5.27	5.27	5.29	5.22	5.29	5.28	5.28
	男子	5.86	5.95	5.77	5.88	5.79	5.82	5.82	5.64	5.82	5.81

- 男女とも  $N = 4000$  である。
- 混合分布としては, かなり「分離」のよい, 良質なものと言えよう。

## シミュレーション： 5種のモデルのパラメータ推定

- 各標本について、全体を1つのクラスとみなす（モデル1）2つのクラスとみなす（モデル2）、3つのクラスとみなす（モデル3）として、MplusのModel: mixtureによる分析を行う。
- さらに下位モデルとして、すべてのクラスの分散が同質（値が等しい）であるとするモデル2\_1とモデル3\_1、分散の値がクラス間で異なるとするモデル2\_2とモデル3\_2を区別した。すなわち、同じ標本を5つの異なるモデルによって分析した。

## シミュレーション：情報量基準

- モデル比較：56400を引いた値である。

標本	基準	モデル					標本	基準	モデル				
		1	2_1	2_2	3_1	3_2			1	2_1	2_2	3_1	3_2
1	AIC	724.9	378.1	363.2	361.3	<b>352.5</b>	6	AIC	378.5	91.8	58.0	57.1	<b>52.9</b>
	BIC	738.9	406.0	<b>398.1</b>	403.2	408.4		BIC	392.5	119.7	<b>93.0</b>	99.1	108.8
	Adjusted-BIC	732.5	393.3	<b>382.2</b>	384.1	383.0		Adjusted-BIC	386.1	107.0	<b>77.1</b>	80.0	83.4
2	AIC	578.1	213.2	<b>180.2</b>	183.8	182.0	7	AIC	454.9	121.7	111.2	109.9	<b>104.3</b>
	BIC	592.1	241.2	<b>215.1</b>	225.7	237.9		BIC	468.9	149.6	<b>146.2</b>	151.9	160.2
	Adjusted-BIC	585.7	228.5	<b>199.3</b>	206.6	212.5		Adjusted-BIC	462.5	136.9	<b>130.3</b>	132.8	134.8
3	AIC	547.9	143.0	119.3	119.5	<b>117.8</b>	8	AIC	428.6	89.3	81.5	83.6	<b>78.1</b>
	BIC	561.8	170.9	<b>154.2</b>	161.4	173.7		BIC	442.5	117.3	<b>116.4</b>	125.5	134.0
	Adjusted-BIC	555.5	158.2	<b>138.3</b>	142.3	148.3		Adjusted-BIC	436.2	104.6	<b>100.5</b>	106.5	108.6
4	AIC	598.4	274.8	265.2	264.1	<b>258.7</b>	9	AIC	439.5	130.2	90.4	<b>84.5</b>	85.0
	BIC	612.3	302.8	<b>300.1</b>	306.0	314.6		BIC	453.5	158.1	<b>125.3</b>	126.4	140.9
	Adjusted-BIC	606.0	290.1	<b>284.2</b>	286.9	289.1		Adjusted-BIC	447.2	145.4	109.5	<b>107.4</b>	115.5
5	AIC	485.2	136.5	125.1	123.3	<b>122.9</b>	10	AIC	395.5	15.3	<b>-0.3</b>	-0.2	-0.1
	BIC	499.2	164.5	<b>160.1</b>	165.2	178.8		BIC	409.5	43.3	<b>34.7</b>	48.7	55.8
	Adjusted-BIC	492.9	151.8	<b>144.2</b>	146.2	153.4		Adjusted-BIC	403.2	30.6	<b>18.8</b>	26.4	30.3

## シミュレーション1：情報量基準の解釈

- それぞれのモデルの自由パラメータの数は、モデル1から順に、2, 4, 5, 6, 8 となっている。
- 情報量基準のうち、BICは、モデル2\_2が最善であることを示唆している。他方、AICは、モデル3、あるいはもっと多くのクラス数のモデルを要求しているようである。
- 男女の身長は、平均値のみならず、標準偏差にも差があったから、BICが正しいモデルを示唆していると言える。

## シミュレーション：2つの分布の特徴

		標 本									
		1	2	3	4	5	6	7	8	9	10
実際の2群の統計測度（構成割合は0.5）											
平均値	女子	157.7	157.7	157.8	157.8	157.7	157.9	157.8	157.8	157.9	157.8
	男子	170.7	170.5	170.7	170.7	170.6	170.5	170.6	170.6	170.6	170.5
標準偏差	女子	5.38	5.30	5.26	5.27	5.27	5.29	5.22	5.29	5.28	5.28
	男子	5.86	5.95	5.77	5.88	5.79	5.82	5.82	5.64	5.82	5.81
Mixture model の解											
平均値	Class_1	156.9	156.8	156.6	157.5	157.2	156.6	157.4	157.2	156.5	157.1
	Class_2	169.9	169.8	169.3	170.2	169.9	169.0	170.0	169.9	168.9	169.8
標準偏差	Class_1	4.99	4.74	4.69	5.10	4.98	4.70	5.02	5.02	4.62	4.84
	Class_2	6.18	6.09	6.45	6.15	6.07	6.52	6.11	6.00	6.54	6.06
Class_1 の構成割合		0.439	0.428	0.410	0.467	0.451	0.384	0.459	0.447	0.379	0.446

- どの標本でも、Class\_1（身長の低いクラス）の構成割合が小さい。
- それに対応して、どちらのClassも平均値が小さく、散布度の差が拡大している。

## 解の尤度は、本来の構成割合の尤度より常に高い

		標 本									
		1	2	3	4	5	6	7	8	9	10
実際の2群の統計測度 (構成割合は0.5)											
平均値	女子	157.7	157.7	157.8	157.8	157.7	157.9	157.8	157.8	157.9	157.8
	男子	170.7	170.5	170.7	170.7	170.6	170.5	170.6	170.6	170.6	170.5
標準偏差	女子	5.38	5.30	5.26	5.27	5.27	5.29	5.22	5.29	5.28	5.28
	男子	5.86	5.95	5.77	5.88	5.79	5.82	5.82	5.64	5.82	5.81
Mixture model の解											
平均値	Class_1	156.9	156.8	156.6	157.5	157.2	156.6	157.4	157.2	156.5	157.1
	Class_2	169.9	169.8	169.3	170.2	169.9	169.0	170.0	169.9	168.9	169.8
標準偏差	Class_1	4.99	4.74	4.69	5.10	4.98	4.70	5.02	5.02	4.62	4.84
	Class_2	6.18	6.09	6.45	6.15	6.07	6.52	6.11	6.00	6.54	6.06
Class_1 の構成割合		0.439	0.428	0.410	0.467	0.451	0.384	0.459	0.447	0.379	0.446
最大対数尤度											
本来の構成割合		-28380.4	-28293.7	-28262.4	-28333.2	-28259.6	-28233.4	-28252.0	-28237.9	-28251.6	-28199.8
最尤解		-28376.6	-28285.1	-28254.6	-28327.6	-28257.6	-28224.0	-28250.6	-28235.7	-28240.2	-28194.9

## なぜ、「正しい解」が出ないのか？

- 最尤解が、想定したものと異なる理由として、性別の身長分布の正規分布とのずれ、測定値の刻みが粗すぎることで、などが考えられたが、いずれも適合しなかった。
- しかしながら、意外なところ (software 実行時の設定ミス) から、(このシミュレーションデータのような構造をもつデータに限り) 正解に近いものを得る方法が見出された。
- 2つの標本を2つの変数とみなして分析を実行することである。
- 標本1と2の同時分析の結果を示す。他の組み合わせでも、同様の結果であった。

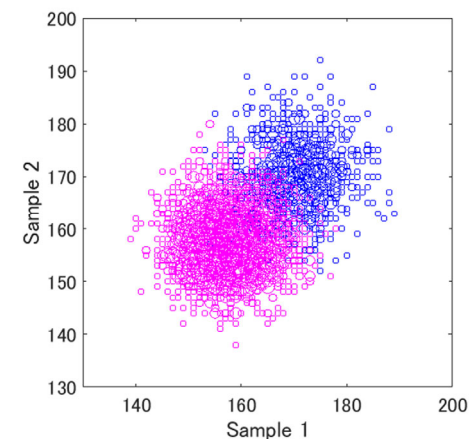
## 同時分析の結果はデータ生成の設定に近い

- 次は、標本1と標本2を同時に分析した結果 (2変数と表示) を、他の条件と比較したものである。
- 統計測度も対数尤度も設定に近い。構成割合は当然どちらの標本でも同じである。

	Class_1		Class_2		構成割合	対数尤度
	平均値	標準偏差	平均値	標準偏差		
標本1						
設定	157.7	5.38	170.7	5.86	0.500	-28380.4
単独	156.9	4.99	169.9	6.18	0.439	-28376.6
2変数	<b>157.7</b>	<b>5.36</b>	<b>170.6</b>	<b>5.94</b>	<b>0.495</b>	<b>-28380.4</b>
標本2						
設定	157.7	5.30	170.5	5.95	0.500	-28293.7
単独	156.6	4.69	169.3	6.45	0.410	-28285.1
2変数	<b>157.7</b>	<b>5.18</b>	<b>170.5</b>	<b>6.01</b>	<b>0.495</b>	<b>-28291.6</b>

## 同時分析とは？

- 2つの標本は独立であり、「対応のあるデータ」ではないから、本来こういう分析はできない。
- しかし、これらの標本では、前の4000行が女子、後の4000行が男子と別れていたため、右の図のように級内では独立、級間には変動があるというデータになった。

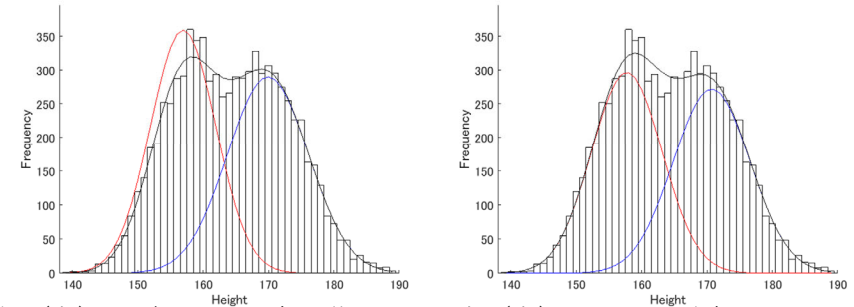


## 何をもって解とすべきなのか

- ここで、1つの変数を対象に、有限混合モデルの解を求めるならば、「正しい」解は最尤解というべきであろう。
- しかしながら、このシミュレーションでは、それは「正解」ではなく、むしろ2つの標本を組み合わせた（本来、正しい分析方法とは言えない）2変数のときに、設定された母集団の比率に近い構成割合が得られた。
- もう1つ、組み合わせた解の方が、標準誤差が小さい（単独の解の半分程度）ことも、考慮する必要がある。

## 2つの解に見分けがつくほどの差があるのか

- 混合分布（標本1）での2つの方法による密度関数の比較



最尤解（左）とデータから直接導かれる分布（右）。ほとんど変わらないように見えるが2つのグループの比率の差は大きい（0.439と0.500）。

## モデル2\_2の最尤解には特異点がある

- このようなことが起こる理由としては、サンプルサイズが500では不十分であること等も考えられるが、実は、最尤解に限って、モデル2\_2に不都合な特異点があることが知られている。
- すなわち、任意の個体のみをクラス1、残りのすべての個体をクラス2とし、クラス1の平均値はその個体の座標、分散は0すると、その個体の対数尤度は無限大になる（正規分布は1点に収斂し、確率密度はその点で $\infty$ 、他のすべての点で0となる）。もう1つのクラスの分布は、分散を正の値にしておけば、対数尤度関数は無限大となり、それが最適解になる。
- このことが、モデル2\_2の「偏り」や標準誤差が大きいことと関連しているかどうかは、今のところわからないというしかない（報告者の能力では）

## Mixture modeling をどう活用するか？

- かなり条件の良いデータでのシミュレーションではあるが、潜在グループの分散が異なる場合、それらの分布を最尤法のプログラムで再現することは困難であることが明らかになった。
- その理由は、現在のところ明らかではないが、こうした結果は、実用上は極めて不都合なことであるが、データ分析のみならず、結果の活用にあたっては、認識しておかなければいけないことであろう。

## さらなる疑問と回答の試み

最近の臨床診断研究の紹介

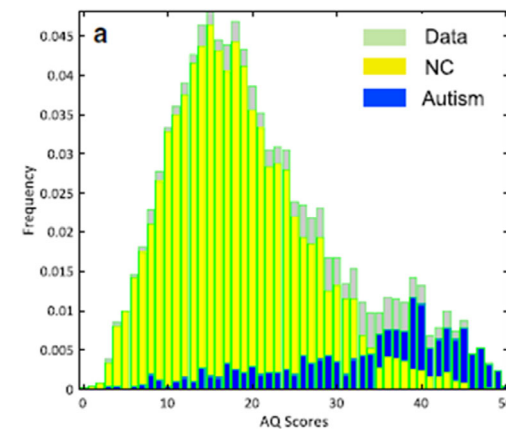
### 用いられた検査

- Baron-Cohen et al. (2001) による 50 項目の AQ scale.
- 高機能自閉症診断のための self-rating 項目
- スコアは, 0 ~ 50点。本来, 5つの sub-scale からなるが, この研究では全項目の合計点 (sum score) が用いられている。

### 2つの疑問とそれへの対応

- 潜在グループごとの得点は, 正規分布であると仮定してよいのか?
- (2つの) グループの混合比率と異なるデータ (たとえば, 臨床群の割合を多くしたデータ) で, 正確な推定はできるのか?
- Abu-Akel, A., Allison, C., Baron-Cohen, S., & Heinke, D. (2019). The distribution of autistic traits across the autism spectrum: evidence for discontinuous dimensional subpopulations underlying the autism continuum. *Molecular Autism*, <https://doi.org/10.1186/s13229-019-0275-3>.

臨床群の得点が低い方に伸びる傾向は, この例でもある



## Weibull 分布の利用

- 障がい者の負に偏った分布と健常者のせいに偏った分布を混合するために、基本になる分布として Weibull 分布を採用
- このようになる初等的な分布はほとんどない（ $\beta$ 分布？）。
- Weibull 分布とは？

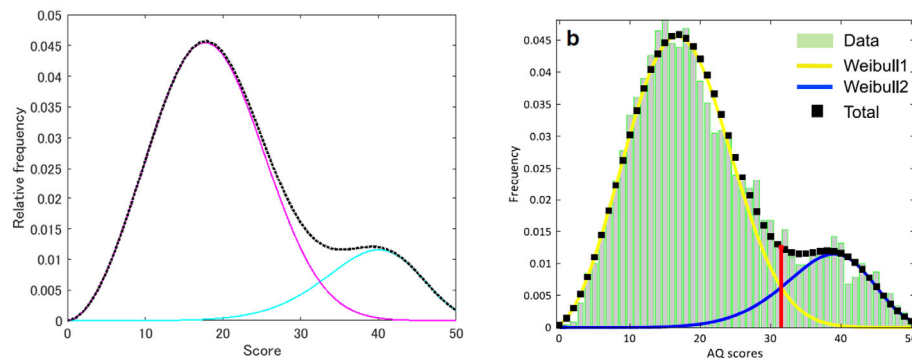
- $$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp[-(x/\lambda)^k], & x \geq 0 \\ 0, & x < 0. \end{cases}$$
- ここで、 $k > 0$  は形状パラメータ（shape parameter）、 $\lambda > 0$  は尺度パラメータ（scale parameter）である。

## Weibull 分布の特徴

- Weibull 分布の平均値、分散、歪度は次のようになる。
- $\mu = \lambda \Gamma(1 + 1/k)$
- $\sigma^2 = \lambda^2 \Gamma(1 + 2/k) - \mu^2$
- $\varpi = [\lambda^3 \Gamma(1 + 3/k) - 3\mu\sigma^2 - \mu^3]/\sigma^3$
- 形状パラメータ  $k$  の値により、正負の偏りを作り出せる。
- データによる推定値は、右の表のようになったという（ $\mu$ 、 $\sigma$ 、skewness は報告者の計算）。

	Group 1	Group 2
$k$	2.9	7.1
$\lambda$	20.5	40.9
$\mu$	18.3	38.3
$\sigma$	6.87	6.36
skewness	0.21	-0.47

## データへの当てはまりはますます良い



## 健常者と障害者の比率を 59 : 1 とする

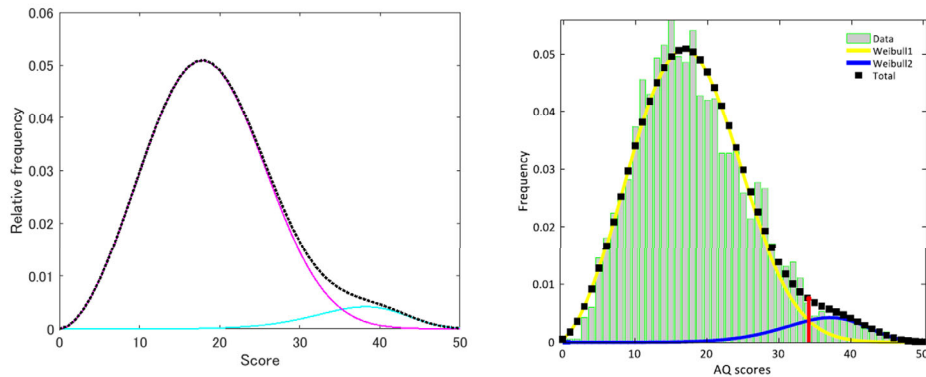
	Group 1	Group 2
$k$	2.9	7.1
$\lambda$	20.5	40.9
$\mu$	18.3	38.3
$\sigma$	6.87	6.36
skewness	0.21	-0.47

→

	Group 1	Group 2
$k$	2.8	7.3
$\lambda$	20.7	39.0
$\mu$	18.5	36.6
$\sigma$	7.04	5.91
skewness	0.22	-0.49

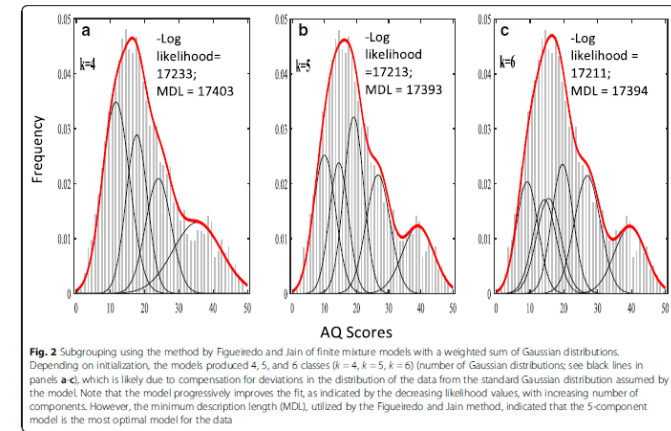


## サンプリング後の分布



Cut point がやや高くなるが、大筋で結果に違いはない。

## 正規分布では少なくとも 4-class 解が必要



**Fig. 2** Subgrouping using the method by Figueiredo and Jain of finite mixture models with a weighted sum of Gaussian distributions. Depending on initialization, the models produced 4, 5, and 6 classes ( $k=4, k=5, k=6$ ) (number of Gaussian distributions; see black lines in panels a-c), which is likely due to compensation for deviations in the distribution of the data from the standard Gaussian distribution assumed by the model. Note that the model progressively improves the fit, as indicated by the decreasing likelihood values, with increasing number of components. However, the minimum description length (MDL), utilized by the Figueiredo and Jain method, indicated that the 5-component model is the most optimal model for the data.

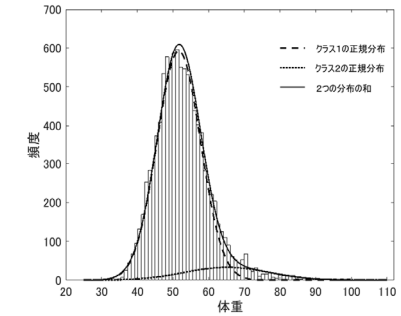
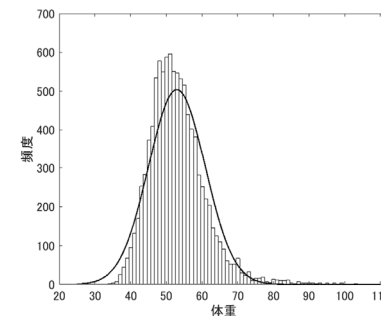
## なお残る疑問

- AQ は自然な原点のある比率尺度なのか？ Weibull 分布を用いる限り、避けて通れない問題。以下の疑問にもつながる。
- 本来多次元である 50項目を単純加算したスコアは、分布の形状を推定するに足るほどの量的特性をもつのか。
- 同様に、臨床群を 1つの群と考えてよいのか？

## 全ての偏った分布が混合分布とは限らない

- 「17歳の女子は、体重の分布の異なる 2つのグループからなる」???

**身長が正規分布するとすれば、体重が正規分布するはずがない。 → BMIの定義を想起！**



## 結論：Mixture modeling の適用について

- ここで紹介したような、極めて初歩的レベルにおいても、mixture modeling には方法自体の問題点とともに、実用上の落とし穴も多いように思われる。
- 量的（質的）データから、カテゴリカルな結論を導く方法の魅力は大きいですが、適用には慎重さが必要である。
- さらに多様な方法群（混合因子分析など）に関しては、今後の研究が待たれる部分もあるが、さらなる分析経験の蓄積も必要であろう。
  
- しかし、これらのことは「直接観測不可能」な潜在変数に依拠する研究全般の危険性を示唆してはいないだろうか？

ご清聴、ありがとうございました



# 縦断データ解析における SEMの応用

お茶の水女子大学 伊藤大幸

## 構成

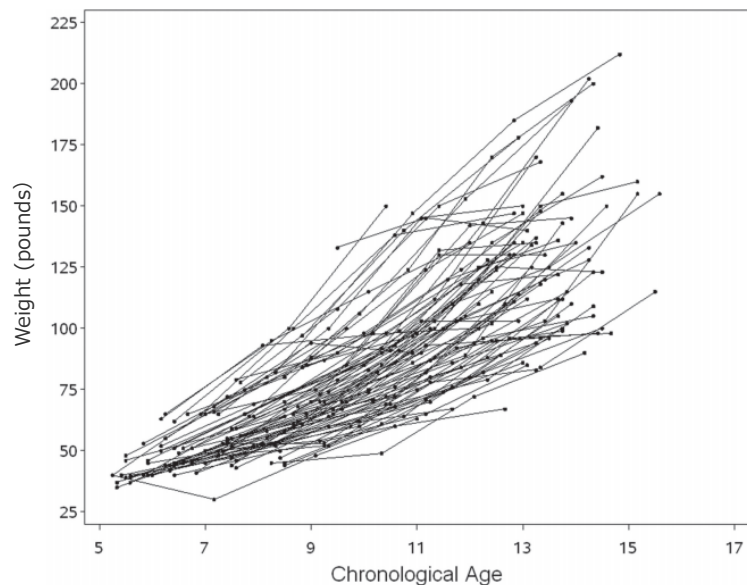
- ◆ 縦断研究と成長モデル
- ◆ 分析例：不登校に至る心理社会的適応の軌跡

## 構成

- ◆ 縦断研究と成長モデル
- ◆ 分析例：不登校に至る心理社会的適応の軌跡

## 縦断研究の意義

1. 年齢・時期・コホートの効果を分離できる
  - 反復横断データでも代替可能 Age-Period-Cohort分析
2. 因果関係に関する手がかりを得られる
  - 交絡の問題は解消されない 交差遅延モデルなど
  - 実験や準実験など、より適したデザインがある場合も
3. 個人内変動の軌跡とその個人差を定量化できる
  - 縦断研究でなければ検証困難 成長モデルなど



個人内変化の軌跡にはある程度共通したパターンがあると同時に個人差もある

## 成長モデルの歴史

- ◆ 成長モデルの起源は、異なる栄養条件下での豚の体重増加に関する Wishart (1938) の分析に遡る
  - 各個体の体重を y 軸、時間を x 軸に取り、データをプロット
  - 各個体の成長軌跡を表す二次関数（曲線）のパラメータを推定
    - ✓ 切片、線形変化（速度）、二次変化（加速度）
  - 栄養条件を独立変数とした分散分析により、各パラメータの有意差を見出し、成長軌跡の違いを示した

## 成長モデルの歴史

- ◆ Wishart が用いた方法は、現代の成長モデルの中核的な側面の理解の基盤を与える
  1. 観察された個人の変化の軌跡は少数のパラメータからなる数学的関数によって近似できる
  2. 関数のパラメータの変動は変化のプロセスにおける個人差に関する情報を反映する
  3. 成長パラメータの変動と他の変数との関連は、その個人差の決定因や結果に関する情報をもたらす

## 成長モデルの歴史

- ◆ 1970年代以降、2つの枠組みの中で、現代的な成長モデルが発展を遂げた
  - マルチレベルモデル（混合効果モデル）
    - ✓ Harville, (1977)、Bryk & Raudenbush (1987)
  - 構造方程式モデリング（SEM）
    - ✓ Meredith & Tisak (1984, 1990)、McArdle (1988)

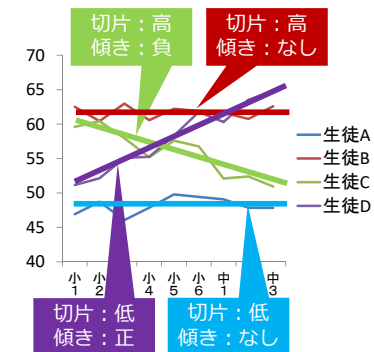
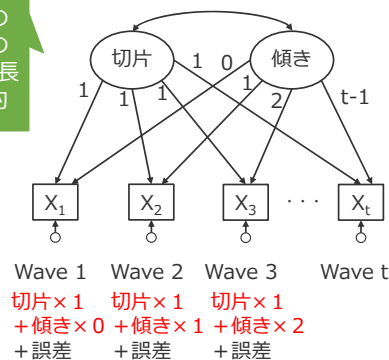
## 2つの枠組みの違い

	マルチレベルモデル	SEM
データ形式	Long形式	Wide形式
測定時点が個人間で異なるデータ	○	△
非線形モデル	○	△
潜在変数の利用	△	○
複数の変数の軌跡	△	○

## SEMによる成長モデルの表現

### ◆ 潜在成長曲線モデル

個人内変化の軌跡を少数の潜在変数（成長因子）で要約



## 成長モデルが扱えるリサーチクエスチョン

1. 標準的な変化の軌跡
2. 変化の軌跡の個人差
  - 例：道徳性の発達の軌跡にはどのようなパターンがあるか
3. 変化の軌跡を予測する要因
  - 例：小児期逆境体験が生涯の精神的健康の軌跡を予測するか
4. 変化の軌跡によって予測される結果
  - 例：心理社会的適応の軌跡によって行動問題の発生が予測されるか
5. 複数の変化の軌跡の関連
  - 例：実行機能と心の理論の発達の発達の連動するか

## 構成

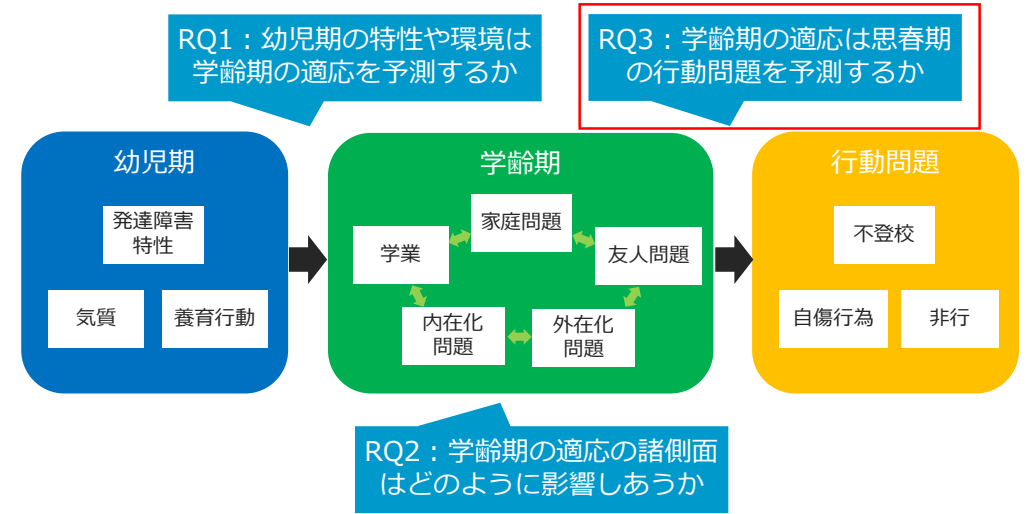
### ◆ 縦断研究と成長モデル

### ◆ 分析例：不登校に至る心理社会的適応の軌跡

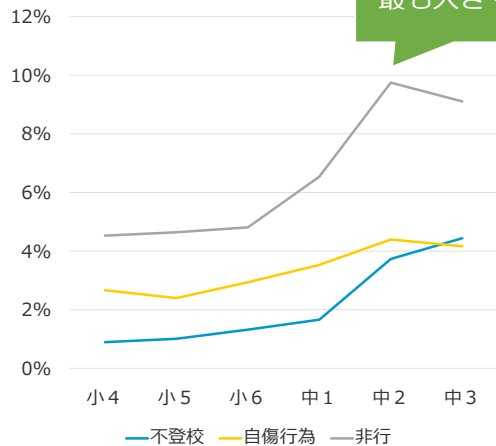
# プロジェクトの紹介

- ◆ 発表者が参画する研究グループによる大規模縦断研究のデータを分析の実例として取り上げる
  - 不登校、いじめ、非行、自傷行為などの情緒・行動問題の発生機序と予防方策の検証を目的とする
  - 2008年から調査協力市の全ての保育所・小中学校で約1万人を対象とする縦断調査を1年間隔で実施
    - ✓ 支援活動の提供により協力関係を構築し、約9割の有効回答率を維持
    - ✓ 児童生徒・保護者・教師の三者から評定を得る
    - ✓ 調査規模・継続期間の両面で国内最大規模の縦断研究

# 研究の枠組み

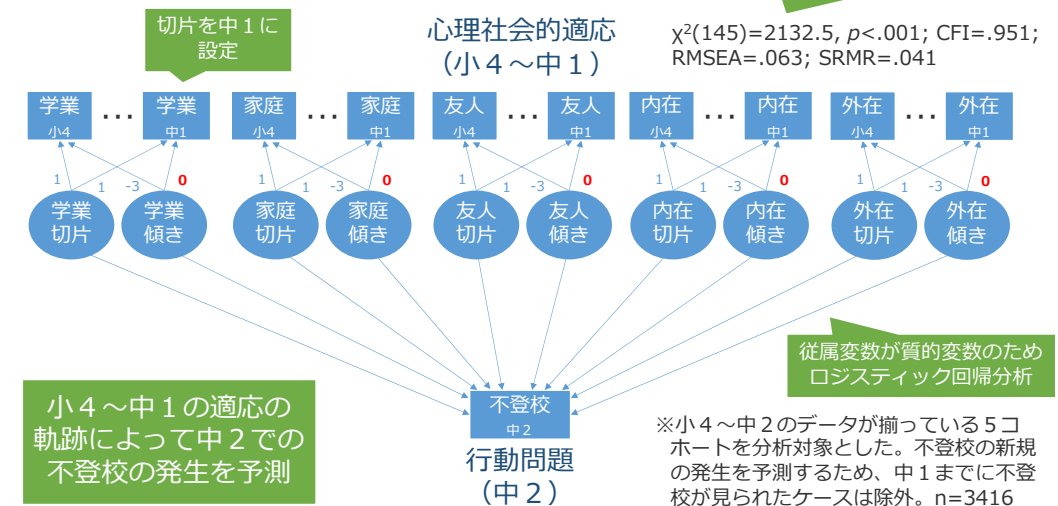


# 行動問題の経験率の推移



どの行動も中2で最も大きく増加

# 成長モデルの指定



# 分析結果

不登校の発生を従属変数とするロジスティック回帰分析の結果

切片のみによる予測

切片&傾きによる予測

	モデル1		モデル2	
	OR	p	OR	p
学業成績 切片	<b>0.77</b>	<b>.036</b>	<b>0.61</b>	<b>.000</b>
家庭問題 切片	<b>1.46</b>	<b>.002</b>	<b>1.34</b>	<b>.030</b>
友人問題 切片	0.86	.342	0.93	.641
内在化問題 切片	<b>1.72</b>	<b>.000</b>	<b>1.64</b>	<b>.001</b>
外在化問題 切片	0.98	.899	0.94	.663
学業成績 傾き			<b>0.54</b>	<b>.000</b>
家庭問題 傾き			1.17	.171
友人問題 傾き			1.11	.383
内在化問題 傾き			1.18	.146
外在化問題 傾き			0.88	.311
Nagelkerke R <sup>2</sup>	.092		.143	

OR：オッズ比  
独立変数が1SD  
上昇したときに  
不登校の発生リスク  
が何倍に上昇するか

・学業成績の水準が低いほど  
・家庭問題の水準が高いほど  
・内在化問題の水準が高いほど  
・学業成績が低下しているほど  
不登校になりやすい

説明率は低い  
→全員共通のモデルで不登校を  
予測することに無理がある？

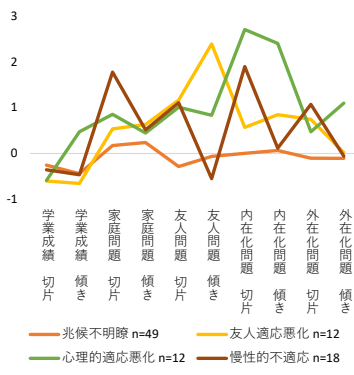
注：ORは各独立変数が1上昇したときのオッズ比を表す。解釈を容易にするため各独立変数はあらかじめ標準化した。モデル1は各要因の切片のみ、モデル2は切片と傾きを独立変数としたモデル。

# 成長混合分布モデル

## ◆ 成長モデルと混合分布モデルを組み合わせた手法

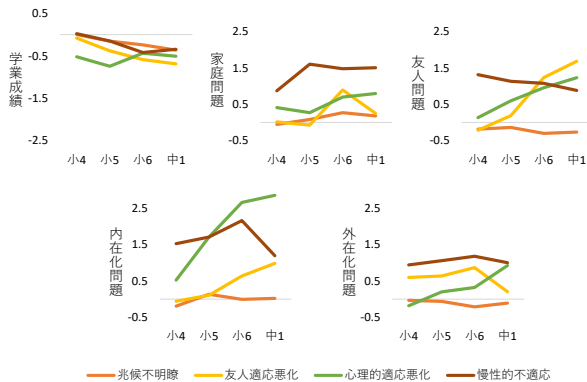
- 縦断的変化の異なる複数の未知の下位集団を同定する
  - ✓ 既知の集団間の差異を検証する多母集団モデルとは異なり、データから事後的に下位集団を同定
- 5つの適応指標における小4～中1の変化の軌跡（切片・傾き）を複数の類型に分けることを試みた
  - ✓ 推定の安定性のため、初めに成長モデルで切片・傾きの因子得点を得て、それを混合分布モデルの推定に使用した
  - ✓ 混合分布モデルには中2で初めて不登校を経験した91名のデータのみを使用

【クラスプロフィール】  
(各群の成長因子の平均値)



適合度、分類確率、人数比、解釈可能性から4クラス解を採用

【各群の小4～中1の適応の軌跡】



不登校に至る心理社会的適応の軌跡には複数の異なるパターンが存在

各クラスを従属変数とするロジスティック回帰分析の結果

「兆候不明瞭群」以外の3群に対しては高い説明力

	兆候不明瞭群				友人適応悪化群				心理的適応悪化群				慢性的不適応群			
	モデル1		モデル2		モデル1		モデル2		モデル1		モデル2		モデル1		モデル2	
	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p
学業成績 切片	0.72	.064	<b>0.59</b>	<b>.005</b>	0.64	.111	<b>0.35</b>	<b>.004</b>	0.62	.213	0.76	.588	1.05	.871	0.91	.766
家庭問題 切片	<b>1.53</b>	<b>.021</b>	1.29	.197	1.35	.338	1.13	.751	0.90	.794	0.99	.978	<b>2.45</b>	<b>.000</b>	<b>2.36</b>	<b>.001</b>
友人問題 切片	<b>0.53</b>	<b>.035</b>	0.61	.105	<b>2.14</b>	<b>.006</b>	0.95	.919	0.45	.087	0.38	.162	0.78	.419	0.99	.967
内在化問題 切片	1.22	.415	1.16	.560	0.73	.396	0.47	.132	<b>6.26</b>	<b>.000</b>	<b>7.39</b>	<b>.000</b>	<b>2.65</b>	<b>.000</b>	<b>2.41</b>	<b>.001</b>
外在化問題 切片	0.78	.285	0.76	.246	1.08	.773	0.97	.921	1.11	.775	1.02	.976	1.30	.270	1.24	.399
学業成績 傾き							<b>0.52</b>	<b>.000</b>			<b>0.28</b>	<b>.006</b>	1.43	.498	0.59	.093
家庭問題 傾き			1.36	.094			1.83	.083			0.68	.337			1.00	.988
友人問題 傾き			1.05	.817			<b>12.35</b>	<b>.000</b>			0.90	.780			<b>0.61</b>	<b>.032</b>
内在化問題 傾き			1.20	.341			1.33	.434			1.44	.250			0.86	.507
外在化問題 傾き			0.74	.098			0.58	.095			<b>3.48</b>	<b>.006</b>			1.04	.891
Nagelkerke R <sup>2</sup>	.035		.093		.087		.538		.368		.543		.299		.351	

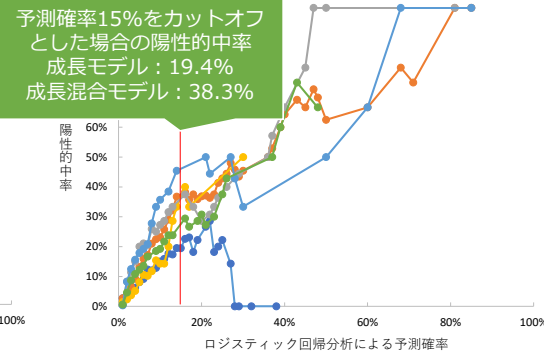
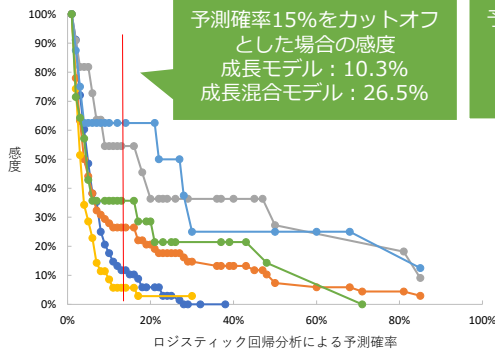
注：ORは各独立変数が1上昇したときのオッズ比を表す。不登校なし群を基準カテゴリとした。解釈を容易にするため各独立変数はあらかじめ標準化した。モデル1は各要因の切片のみ、モデル2は切片と傾きを独立変数としたモデル。

実際に陽性であるケースのうち  
正しく陽性と判定される割合

陽性と判定されたケースのうち  
実際に陽性である割合

【予測確率のカットオフと感度の関連】

【予測確率のカットオフと陽性的中率の関連】



● 成長モデル  
● 成長混合モデル (全体)  
● 成長混合モデル (兆候不明瞭群)  
● 成長混合モデル (友人適応悪化群)  
● 成長混合モデル (心理的適応悪化群)  
● 成長混合モデル (慢性的不適応群)

● 成長モデル  
● 成長混合モデル (全体)  
● 成長混合モデル (兆候不明瞭群)  
● 成長混合モデル (友人適応悪化群)  
● 成長混合モデル (心理的適応悪化群)  
● 成長混合モデル (慢性的不適応群)

## 考察

1. 前年度の状態だけでなく過去数年の軌跡を考慮することが予測の精度を向上させる
  - 児童生徒の適応状態の縦断的なモニタリングが行動問題の発生リスクの推定に貢献
2. 不登校に至るプロセスには複数のパターンがある
  - 全員が同じメカニズムで不登校に陥るわけではない
3. パターンの違いを考慮することで予測精度を高められる
  - 変数指向アプローチと個人指向アプローチの融合

## 論点

- ◆ 唯一の「普遍的」なモデルを想定することの妥当性
  - 大部分の心理学研究では暗黙のうちに全個人に共通のモデルが想定されるが、その想定はほとんど検証されない
    - ✓ 不登校の発生プロセスが児童生徒によって異なることは、学校現場ではほぼ周知の事実
  - 個人によって因果関係や因子構造に差異がある可能性を、もっと積極的に検証する必要があるのでは
  - 縦断データや反復測定データはこうした可能性の検証に貢献
    - ✓ 縦断研究の4つめの意義？