

『パーソナリティ心理学における 統計分析の動向』

荘島宏二郎@指定討論
大学入試センター研究開発部

大久保先生on効果量

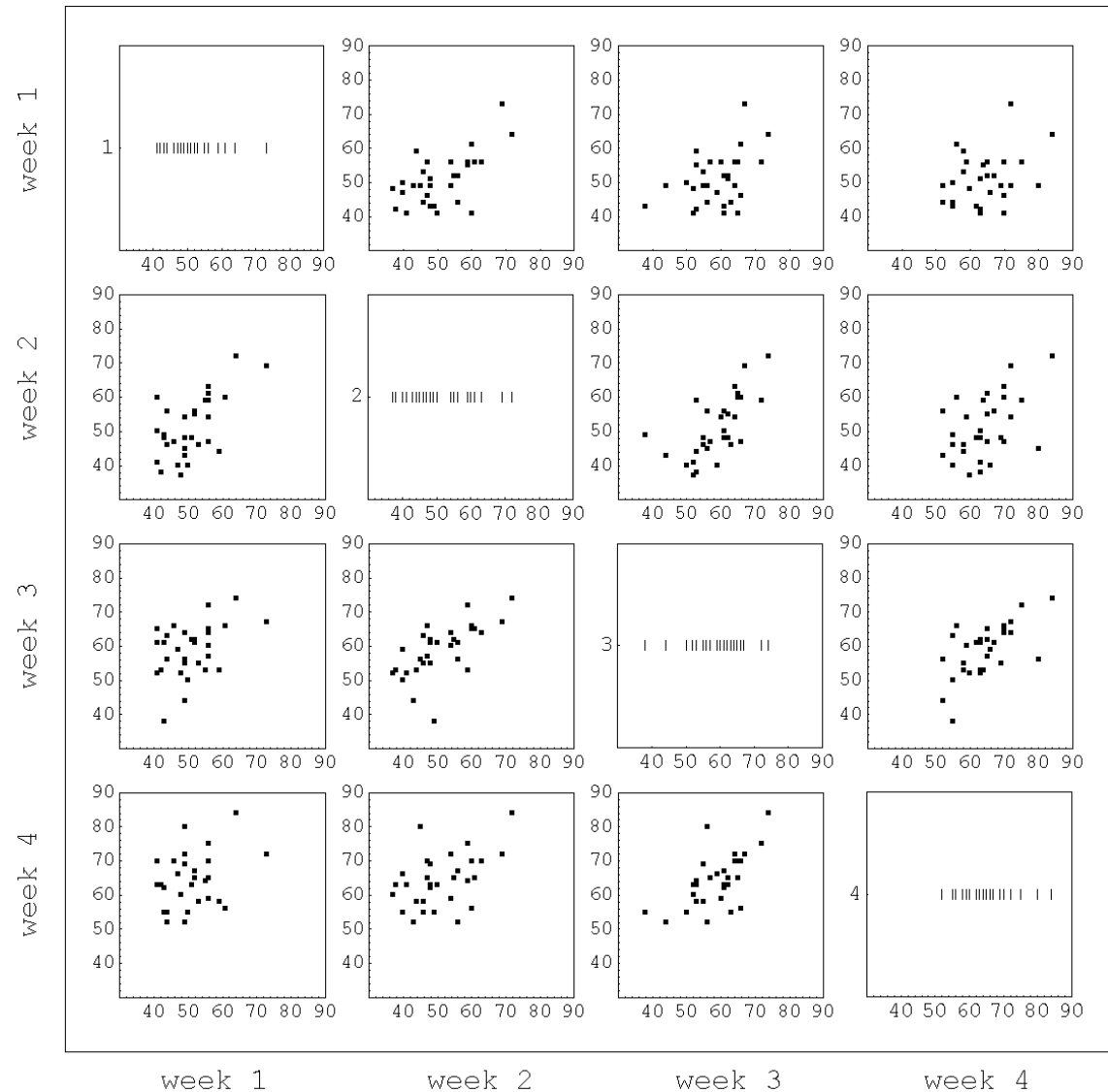
- 南風原(1991) 第33回日本教育心理学会
- 南風原(1995) 教育心理学年報, 34
- 南風原(2002) 心理統計学の基礎 有斐閣
- 大久保・岡田(2012)伝えるための心理統計
勁草書房
 - 2013行動計量学会 杉山明子賞
 - 我が国における効果量の認知・普及に決定的影響
 - メタ分析の普及も加速が予測

p 値 \rightarrow +効果量

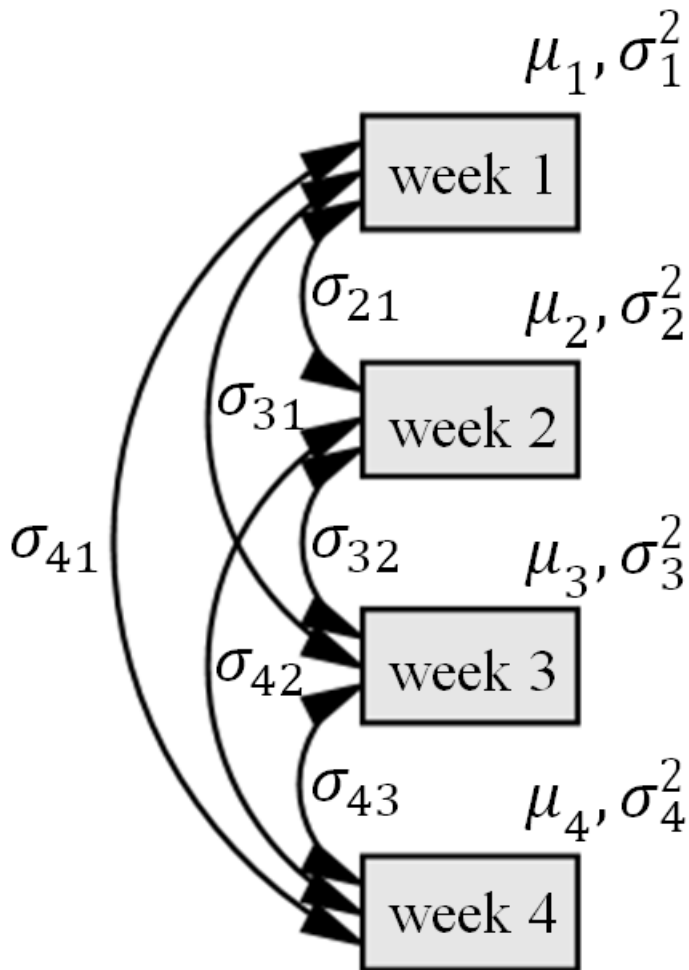
- p 値 = f (効果量, 標本サイズ)
 - 効果量は量の大きさを直接には操作不可
 - 標本サイズは直接操作可能 \leftarrow がんばりどころ
- 「効果量を使う」 = 「人間の直接操作が及ばない量で白黒つけましょうよ」
- 一元的な意思決定 \rightarrow 多元的な意思決定へ
- 全てのデザインにおいて効果量が整備されていないので今後も発展・充実が期待

SEMにおける多元的な意思決定 (被験者内1要因ANOVAを例に)

| | W1 | W2 | W3 | W4 |
|-----|----|----|----|----|
| S1 | 50 | 40 | 50 | 55 |
| S2 | 52 | 55 | 62 | 65 |
| S3 | 42 | 38 | 53 | 63 |
| S4 | 53 | 46 | 55 | 58 |
| S5 | 46 | 47 | 66 | 70 |
| S6 | 49 | 48 | 55 | 69 |
| S7 | 56 | 54 | 60 | 59 |
| S8 | 51 | 48 | 62 | 63 |
| S9 | 56 | 61 | 65 | 65 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| S30 | 44 | 56 | 56 | 52 |



被験者内1要因ANOVAの SEMにおけるパス図的表現



| | 制約 |
|-----|---|
| 制約1 | 等平均制約 $\mu_1 = \mu_2 = \mu_3 = \mu_4$ |
| 制約2 | 等分散制約 $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$ |
| 制約3 | 対角制約 $\sigma_{21} = \sigma_{31} = \sigma_{41} = \sigma_{32} = \sigma_{42} = \sigma_{43} = 0$ |
| 制約4 | 等共分散制約 $\sigma_{21} = \sigma_{31} = \sigma_{41} = \sigma_{32} = \sigma_{42} = \sigma_{43}$ |
| 制約5 | 等時点間等共分散制約 $\sigma_{21} = \sigma_{32} = \sigma_{43}, \sigma_{31} = \sigma_{42}$ |

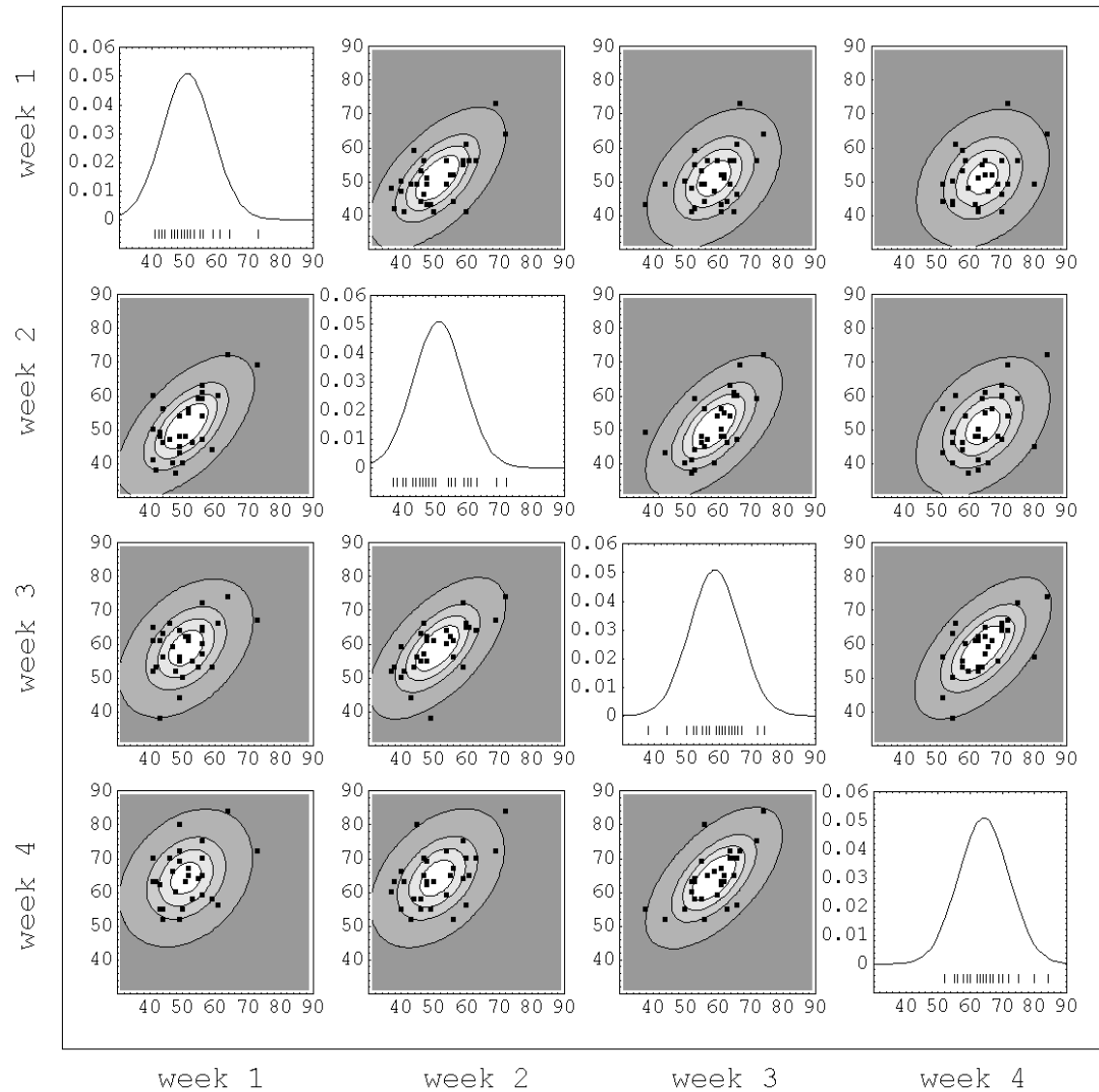
| モデル | 制約 | N_p^* | χ^2 値 | df | p値 | CFI | RMSEA | (HI 90) | AIC | BCC |
|-----------------------------|-----------------------|---------|------------|----|-------|-------|-------|---------|------|------|
| A 飽和 | | 14 | 0.0 | 0 | | 1.000 | | | 28.0 | 33.8 |
| B 等時 | | | | | | | | | | 27.6 |
| C 等共 | | | | | | | | | | 30.1 |
| D 対角 | | | | | | | | | | 63.4 |
| E 等分散 | | | | | | | | | | 28.7 |
| F トープリッツ | | | | | | | | | | 21.6 |
| G 複合 | | | | | | | | | | 24.2 |
| H 等分散 | | | | | | | | | | 57.4 |
| I 等平均 | | | | | | | | | | 60.5 |
| J トープリッツ | | | | | | | | | | 19.2 |
| K トープリッツ | | | | | | | | | | 40.5 |
| L トープリッツ | | | | | | | | | | 55.1 |
| M トープリッツ+ $(\mu_2 = \mu_3)$ | $2, 5, \mu_2 = \mu_3$ | 7 | 32.0 | 7 | 0.000 | 0.344 | 0.351 | (0.478) | 46.0 | 48.9 |
| N トープリッツ+ $(\mu_2 = \mu_4)$ | $2, 5, \mu_2 = \mu_4$ | 7 | 47.1 | 7 | 0.000 | 0.000 | 0.445 | (0.569) | 61.1 | 64.0 |
| O トープリッツ+ $(\mu_3 = \mu_4)$ | $2, 5, \mu_3 = \mu_4$ | 7 | 18.3 | 7 | 0.011 | 0.704 | 0.236 | (0.371) | 32.3 | 35.2 |

「平均が等しい／等しくない」という仮説検定は「平均が等しいモデルを選択するのか、等しくないモデルを選択するのかに読み替え。

ANOVAの決まりきった制約（等分散性、球面性など）に加えて、いろいろな構造を試せる。

多種多様な適合度指標を参考にしてモデル選択できる。

最終モデル



SEMでは

- 多様な適合度指標で複数モデルを多角的に診断して最終モデルを選択
- 仮説検定にもその流れはやってくるかも？
 - 荘島(2008)統計的仮説検定から多重指標によるモデル選択へ 心理学ワールド42
 - 荘島(2014?)検定、効果量、そして適合度
教育心理学年報(執筆中)
- ただし
 - 確証バイアス
 - メタ分析に使いやすい指標ばかりでない

大久保先生へ

- 我が国の査読誌も効果量に関する何らかの規定を設けていくべきでは？
 - 合わせて査読者の教育も必要？
 - (心理)統計教育において必須項目に？
- 効果量に関する基準が普及してきました・・・
 - 新たなアーティファクト？
 - どのようにお考えですか？
- 国内の実験心理学界の受け止めはいかがでしょうか。様子をお教えてください。

高橋先生on縦断データ解析

- 縦断データ解析は非常に専門性が高い。
 - データ収集(ドロップアウト、消耗率)
 - 測定精度(同一尺度の繰り返し測定は学習効果)
 - 分析力(縦断データ解析は職人芸)
- その中で混合軌跡モデリングに着目
 - 発達の軌跡は一様ではない(平均軌跡の誤謬)
 - 明示的でない複数の(非)線形な軌跡が混合
 - 潜在クラス分析+(非)線形回帰分析

高橋先生のRQ

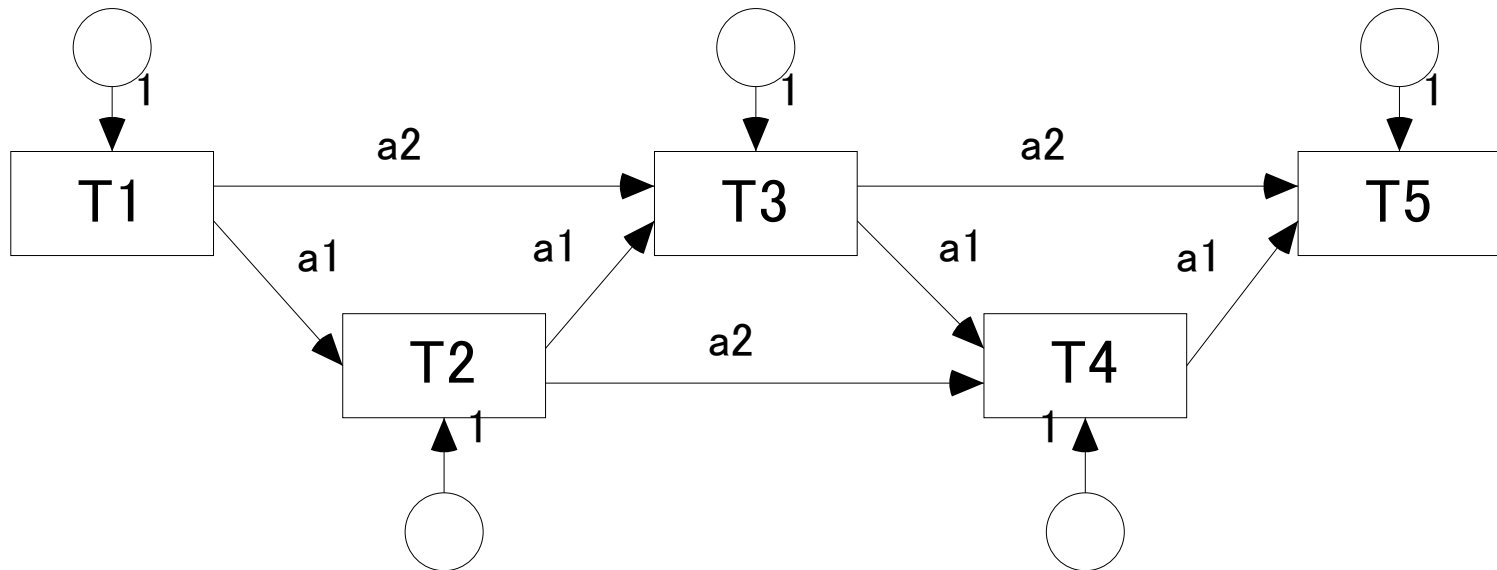
- 発達の様相は一様ではない
 - 認識論、研究態度、信念
- 発達の種類は何パターンなのか
 - 潜在クラス分析(LCA)を適用
- それぞれのパターンはどういう軌跡なのか
 - LCA+非線形回帰分析=混合軌跡モデル
- それぞれのクラスはどのような集団なのか
 - 多項ロジスティック回帰分析
- 目的に応じて適切に統計モデルを選択

混合軌跡モデリングの注意点

- 横軸が時間 T 、縦軸が興味のある Y 変数
 - +ある時点の独立変数 X
- Y on T の非線形回帰分析
- T が直接原因で Y が変化しているのか
- T の推移による X の変化が Y に影響を与える場合
- T の推移($1, \dots, t, \dots T$)
 - X の経時変化($X_1, \dots, X_t, \dots X_{T-1}$)
 - Y の経時変化($Y_2, \dots, Y_{t+1}, \dots Y_T$)
- 高橋さんの文脈では混合軌跡の圧勝
 - この議論は潜在成長にも言える

自己回帰モデル

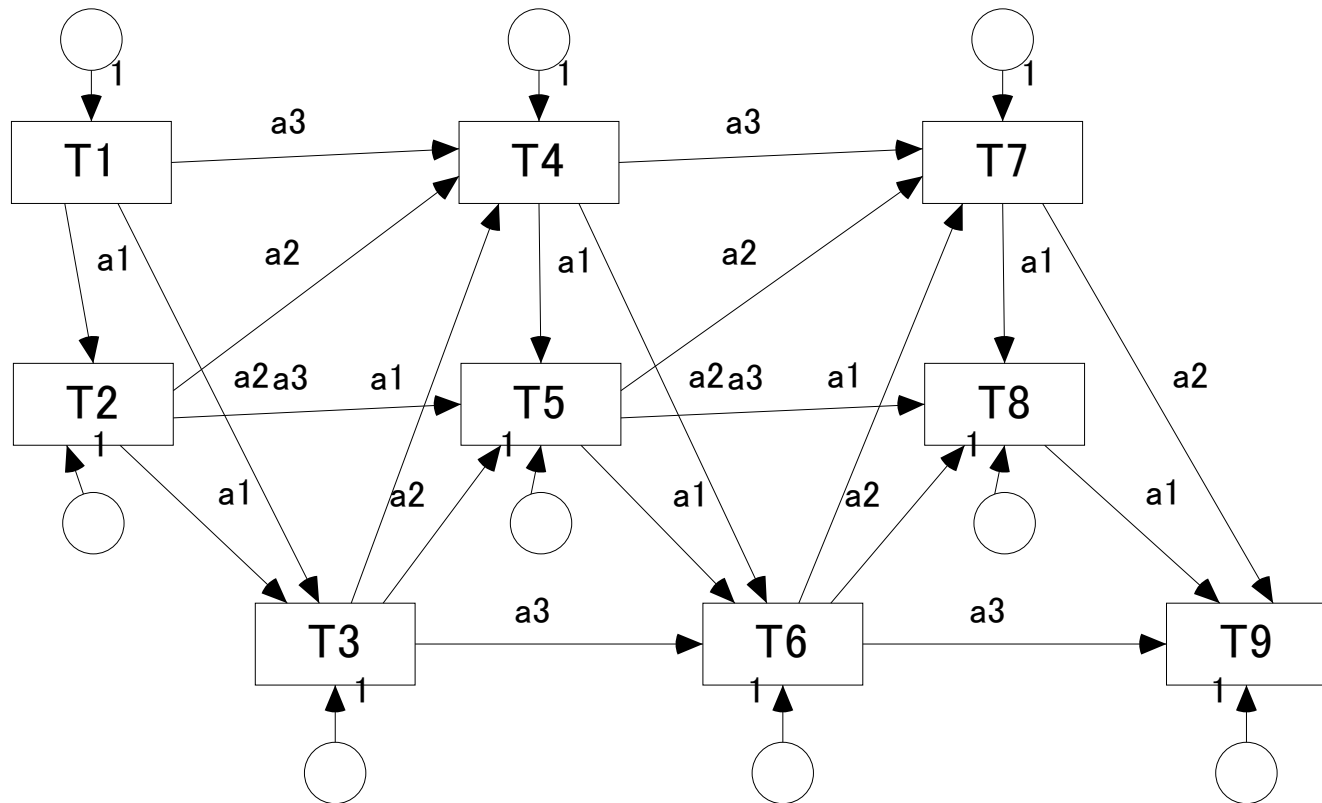
Auto-Regression Model, **AR**モデル



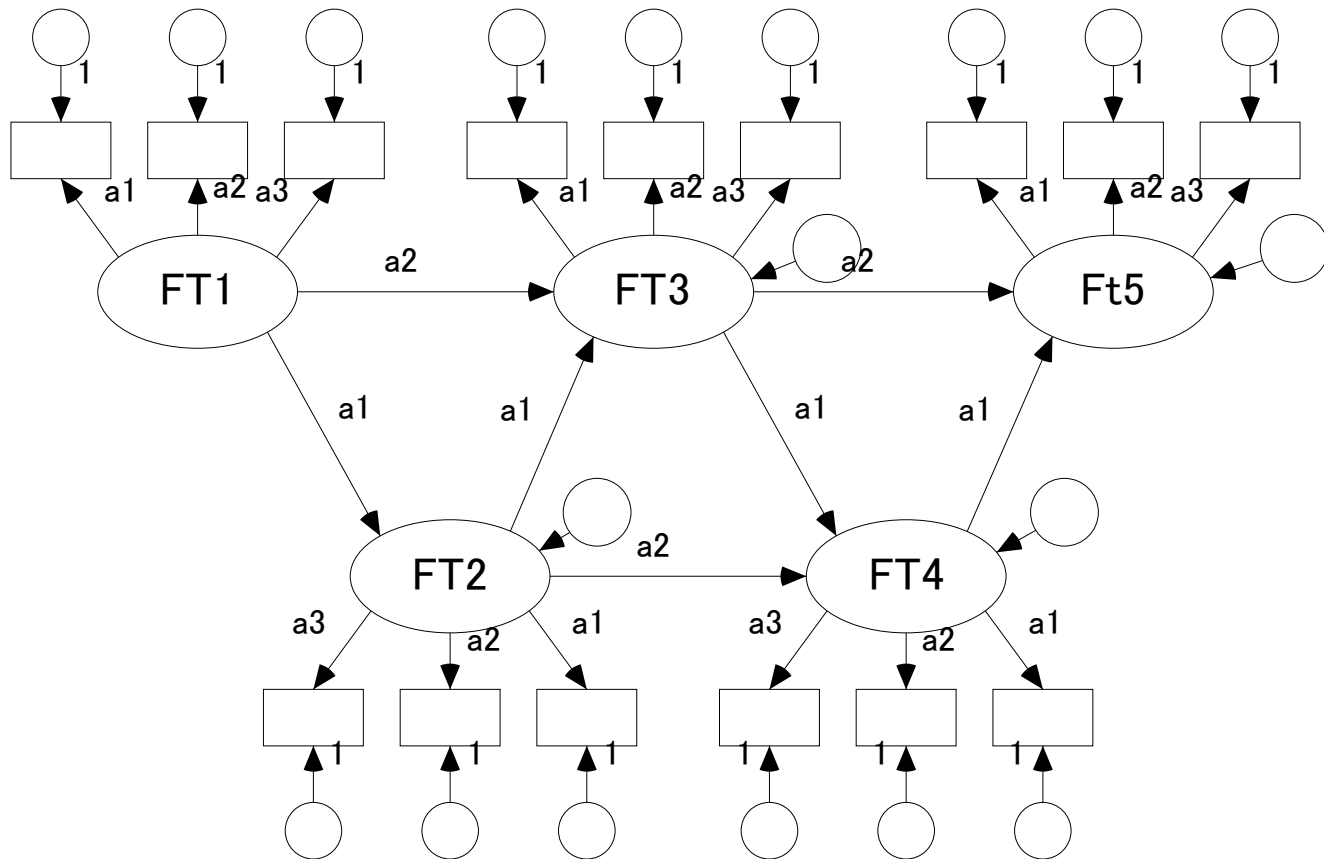
2次の自己回帰モデル

Second-order AR model, AR2

3次の自己回帰モデル, AR3

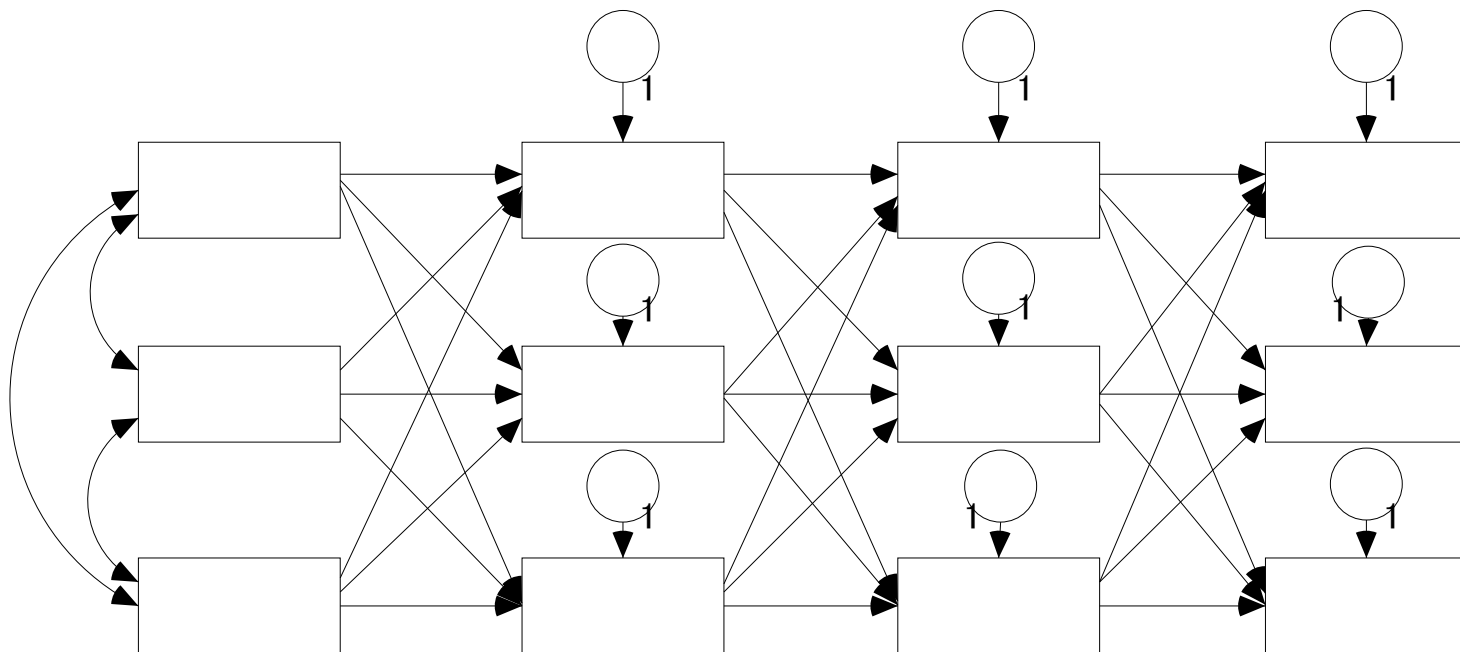


因子のAR2モデル

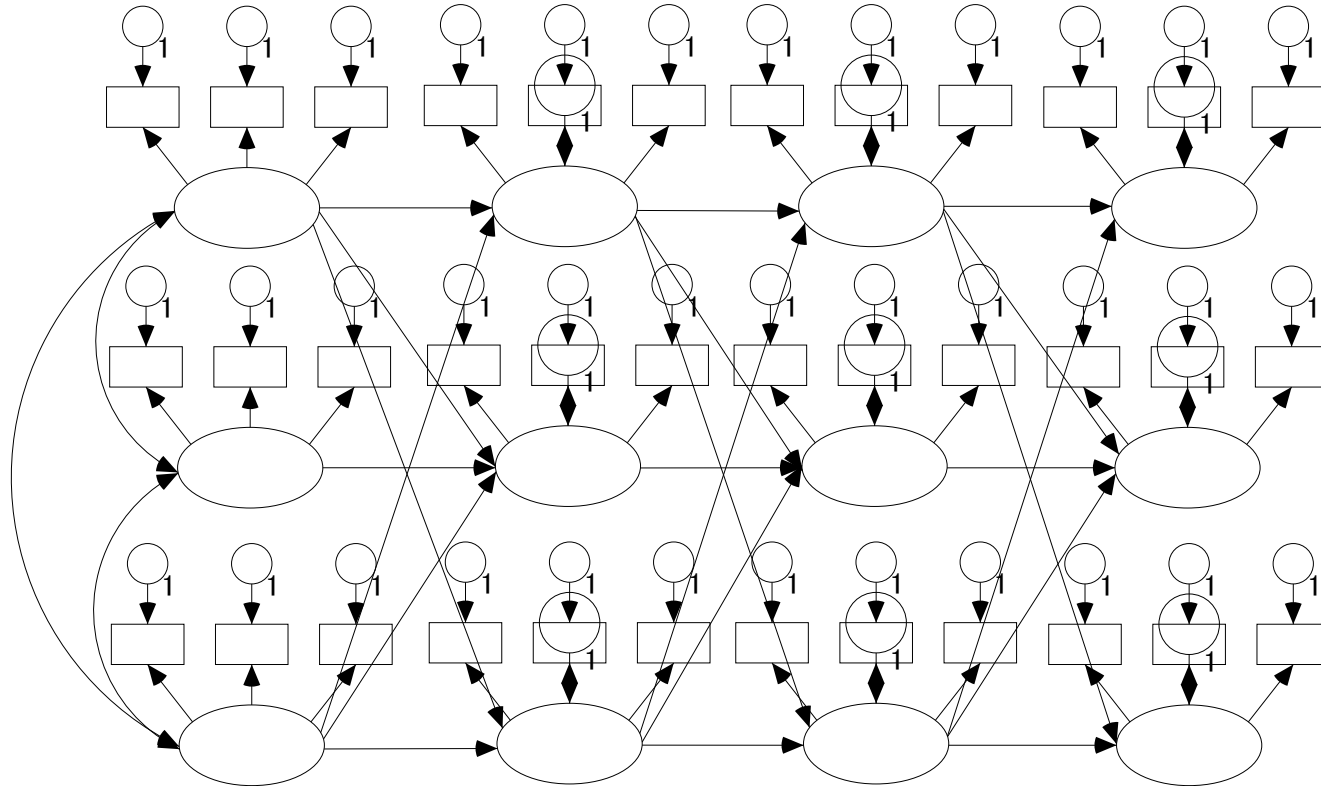


ベクトル自己回帰モデル

Vector Auto-Regression Model, **VAR** Model



因子のVAR(1)



高橋先生へ

- 性差がはっきりしているので男女別（あるいは多母集団）分析しても面白いかも
 - 男女一緒の分析
 - 兄弟の有無の影響が出ない
 - (例) 男女別に分析
 - 男子には兄弟の有無の効果あり
 - 女子には兄弟の有無の効果なし
- 少し社会学的ですが、SES・居住環境・同居人数などの社会・経済・環境変数の影響も気になるところです（今後に期待）

岡田先生onメタ分析

- 山田・井上(2012)以前から取り組んでいることの進取性に驚愕
- 効果量の浸透によってメタ分析も普及が加速
- 学知を総合する上で重要な技術
 - 文献の探索・codingは学生にとってよい訓練
- 一次研究を効率よく集めるためのシステム構築が急務
 - 少なくともお金をかけるところで他国に負けるな

「メタ分析は創造的でない」という批判

- 単に資料を収集し結果を(重み付き)平均しているだけじゃないか、という批判
 - 見当違い
- 効果量の大小がどのような変数(調整変数)によって規定されているか
 - 実験計画的手法で検討
- 矛盾→止揚
 - ある研究では効果あり、ある研究では効果なし
 - 性別が調整変数であると同定
 - 「男性は効果なし／女性は効果あり」→新たな知見

アナログ研究 vs N=1研究

- 臨床研究の場合、多くは臨床群の人に対して介入効果を検証したい
 - しかし倫理(審査)の壁
- アナログ研究
 - 非臨床群に対して研究を行い臨床群に対する効果を推測
 - 倫理的な壁が低い
- N=1研究(一事例研究)
 - 臨床群に対して知見を得る
 - しかし一般化しにくい
- メタ分析はN=1研究の結果の総合方法も用意
 - 研究の個別性が高く適格性基準を満たすことが難しい

今こそ少数事例研究を

- 少数事例研究は面白い
 - 個々の現象に対して深く潜れる
 - 現象に深く潜らなければ、豊かな仮説を引き出すことはできない
- そして、メタ分析へ
 - ベースライン期と介入期を設け各期で複数測定
- 広く浅い大量データを**集め分析する**技術から
狭く深い少量データを**分析し集める**技術へ

岡田先生へ

- 一次研究の適格性基準に査読誌であることとなっていました
 - 公表バイアスの問題
 - 紀要や修論・博論も含めてもよかった？
(収集が大変すぎるのは承知で・・・)
- メタ分析は(特に臨床系の)学生や若手研究者にとって良い訓練になりそうです
 - 大学院生にグループワークさせたら力つきそう
 - 教育実践に取り入れたりしていますか？
- メタ分析の大変さと苦労が伝わってきました
 - 一番楽しかったことは何ですか？