

経時データでの関心事とモデルの誤特定

- Primary interest: 平均構造
 - 治療効果、時期効果、共変量効果
 - モデルの誤特定はなし(できる限り少なくする)
- Secondary interest: 相関構造
 - 結果変数間の相関
 - モデルの誤特定はやむを得ない(できる限り少なくする)

平均パラメータ(平均効果、回帰パラメータ、回帰係数)の推定

- 点推定
 - 平均構造のモデルが正しければ、当てはめた相関構造が真であるかによらずバイアスなし
 - 一般化推定方程式 (Generalized Estimating Equation; GEE) を利用
- 分散(標準誤差)、信頼区間
 - モデル分散(通常分散)
 - 平均パラメータの検定、信頼区間があてにならない
 - ロバスト分散
 - あてはめた相関構造によらず、平均パラメータの検定、信頼区間は妥当

ロバスト分散に基づく推測

- 相関構造のモデル化を誤った場合でも、平均パラメータに対する分散の偏りのない推定量
 - 相関構造の誤特定に対して、頑健な (robust) 分散
- ロバスト分散を用いた検定・信頼区間
 - 当てはめた相関構造が真であるかどうかにかかわらず、妥当
 - 作業相関構造 Working Correlation Structure

White (1982), Liang and Zeger (1986)

エストロゲンパッチデータ

共分散構造		傾き				傾きの差
		エストロゲン群		プラセボ群		
独立	点推定値	-2.49		-1.80		-0.69
	SE	0.16	0.16	0.19	0.23	0.28
等分散・等相関	点推定値	-2.46		-1.77		-0.69
	SE	0.13	0.15	0.16	0.22	0.27
1次の自己回帰	点推定値	-2.49		-1.76		-0.73
	SE	0.21	0.14	0.22	0.25	0.25

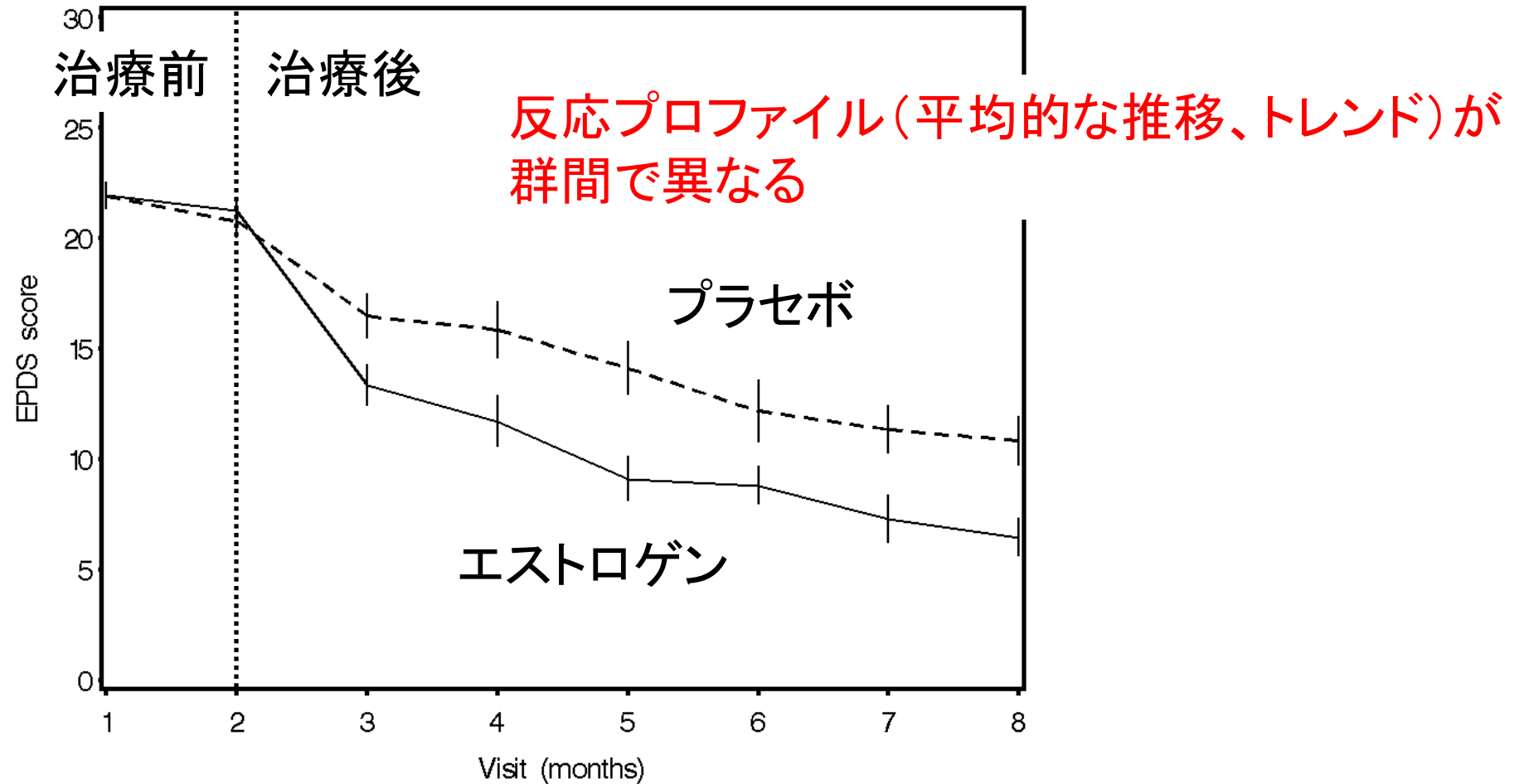
平均構造のモデル化 (1)

- 研究の第一義的な目的
 - 治療効果のバイアスのない推定
 - 平均構造を正しくモデル化
- できる限り仮定の少ないモデル
 - 飽和モデル (Saturated model)
- データのもつ情報量の限界、推定結果の安定性
 - 時点ごとに異なるパラメータ?
 - 経時変化パターン: 線型変化、2次関数的な変化...
- できるだけ簡単なモデル
 - モデルの簡略化 Parsimonious model

平均構造のモデル化 (2)

- ランダム化研究
 - 共変量の影響はまず考えなくてもよい
 - 結果変数の経時変化(時間効果)のモデル化に注意
 - 治療効果、時間効果、
治療と時間の交互作用効果(反応プロファイルの違い)を
モデルに含めるのが常套手段
- 疫学研究
 - 共変量の影響(交絡)を考慮しなければならない
 - 時間効果のモデル化にも注意

治療と時間の交互作用効果



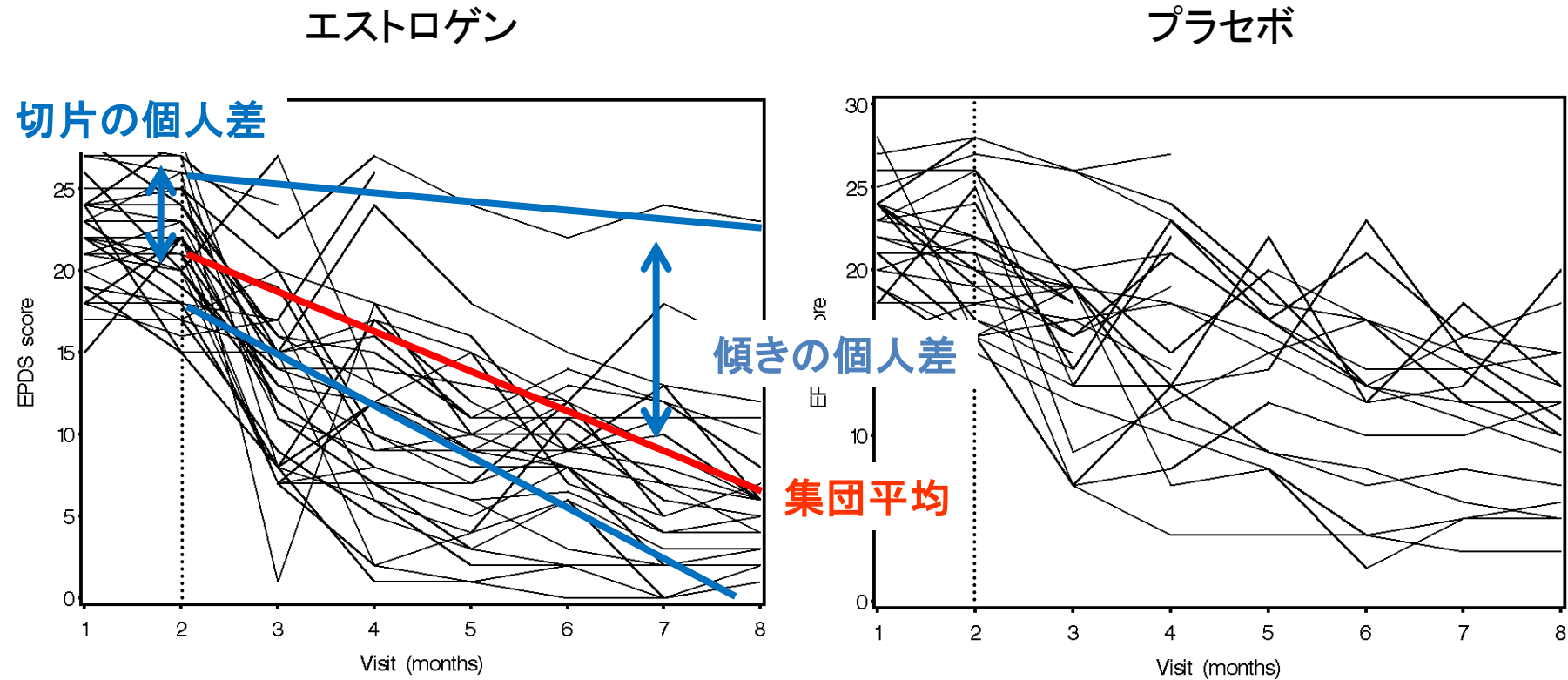
内容

- 経時データとは
- いくつかの例
- 単純な解析
- 2つの解析アプローチ
 - 周辺モデル
 - 混合効果モデル
- 欠測データ(ちょっとだけ)
- まとめ

混合効果モデル

- 個人間の異質性 Heterogeneity
 - 個人間で効果は異なると仮定
- 個人差を表す潜在的な因子
 - 研究で観察されている要因以外に、観察されていない要因が存在
 - この要因は個人間で異なるが個人内では共通であり、よって、複数の結果変数間に相関が生じる(個人特有の効果を与えたもとでは、個人内の結果変数は相関なし)

変化率(傾き)の群間比較



個人特有の効果

- Y_{it} : i 番目の対象者の時点 t (治療開始時をゼロ) における結果変数

$$E(Y_{it}|b_i) = \begin{cases} \beta_0 + \beta_1 t + b_{i0} + b_{i1} t & \text{プラセボ群} \\ \beta_2 + \beta_3 t + b_{i0} + b_{i1} t & \text{エストロゲン群} \end{cases}$$

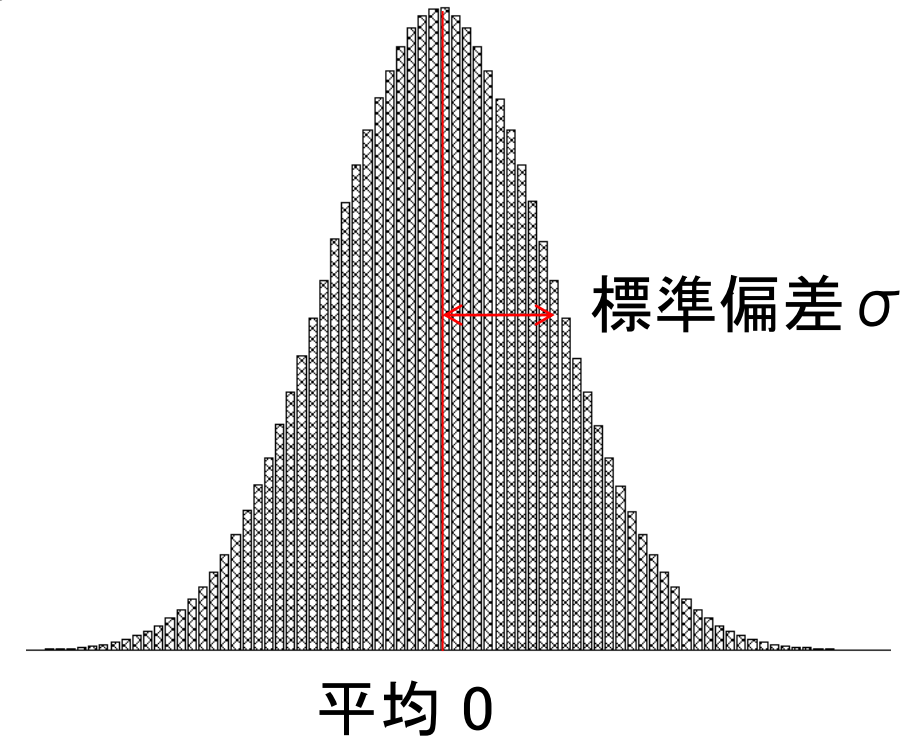
個人効果を与えたもとでの
平均値、期待値

b_{i0} : 切片に関する集団平均からのズレ

b_{i1} : 傾きに関する集団平均からのズレ

個人特有の効果

- 個人特有の効果の推定
 - 通常のサンプルサイズ(時点数)では、仮定をおかずに推定するのは困難
- 仮定
 - 個人特有の効果は、ある一定の大きさでばらついているが、平均を取ればゼロとなる
 - 平均がゼロ、分散(標準偏差)がある値の正規分布からのサンプル



相関の考慮の仕方

- 相関を規定する要因
 - 個人ごとの回帰パラメータ(変量効果)
- 個人特有の効果を与えたもとでは、個人内の結果変数は独立
 - 個人特有の効果で個人内の相関構造を評価

固定効果と変量効果

- 固定効果（母数効果） Fixed Effect
 - 因子の水準に普遍性があり、効果自体の推定に関心がある因子
 - 治療変数、時期変数、共変量...
- 変量効果（ランダム効果） Random Effect
 - 因子の水準はある集団からのサンプルと考えられ、効果自体の推定ではなくそのバラツキの大きさに関心がある因子
 - 測定単位（個人、クラスターなど）に特有の効果

多施設共同研究における施設の取り扱い

- 関心が試験の対象となった施設のみにある場合
→ 固定効果
- 試験が実施可能であった全ての施設からランダムに施設が選択されたとし、想定する母集団に対して推論を行いたい場合
→ 変量効果

混合効果モデル

- 固定効果と変量効果の両方を含んだ回帰モデル
 - 線型混合効果モデル Linear Mixed Effects Model
 - 一般化線型混合効果モデル Generalized Linear Mixed Effects Model
 - 非線型混合効果モデル Non-Linear Mixed Effects Model

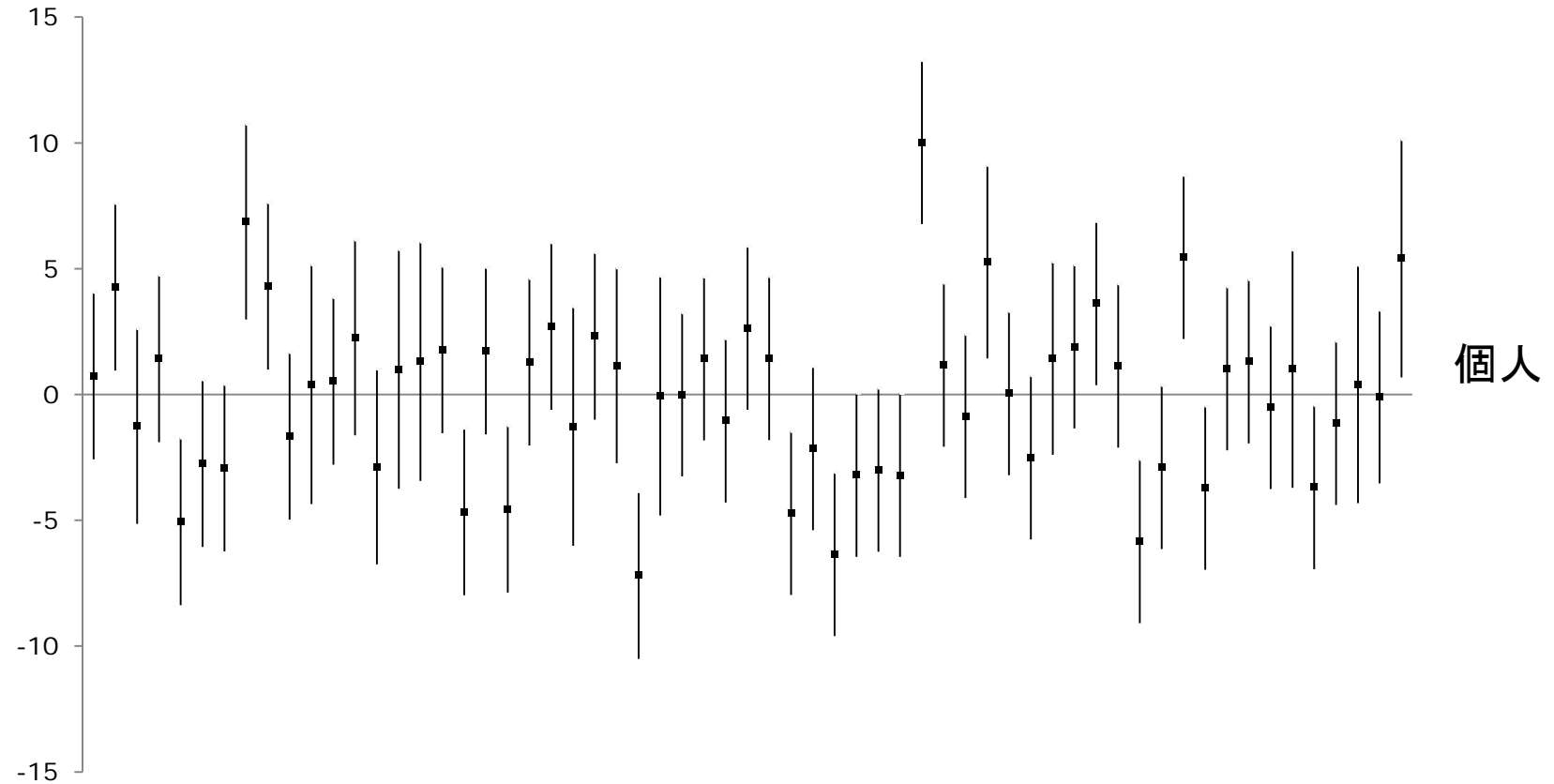
エストロゲンパッチデータ

切片と傾きを変量効果	傾き		傾きの差
	エストロゲン群	プラセボ群	
点推定値	-1.90	-1.53	-0.37
SE	0.18	0.22	0.29

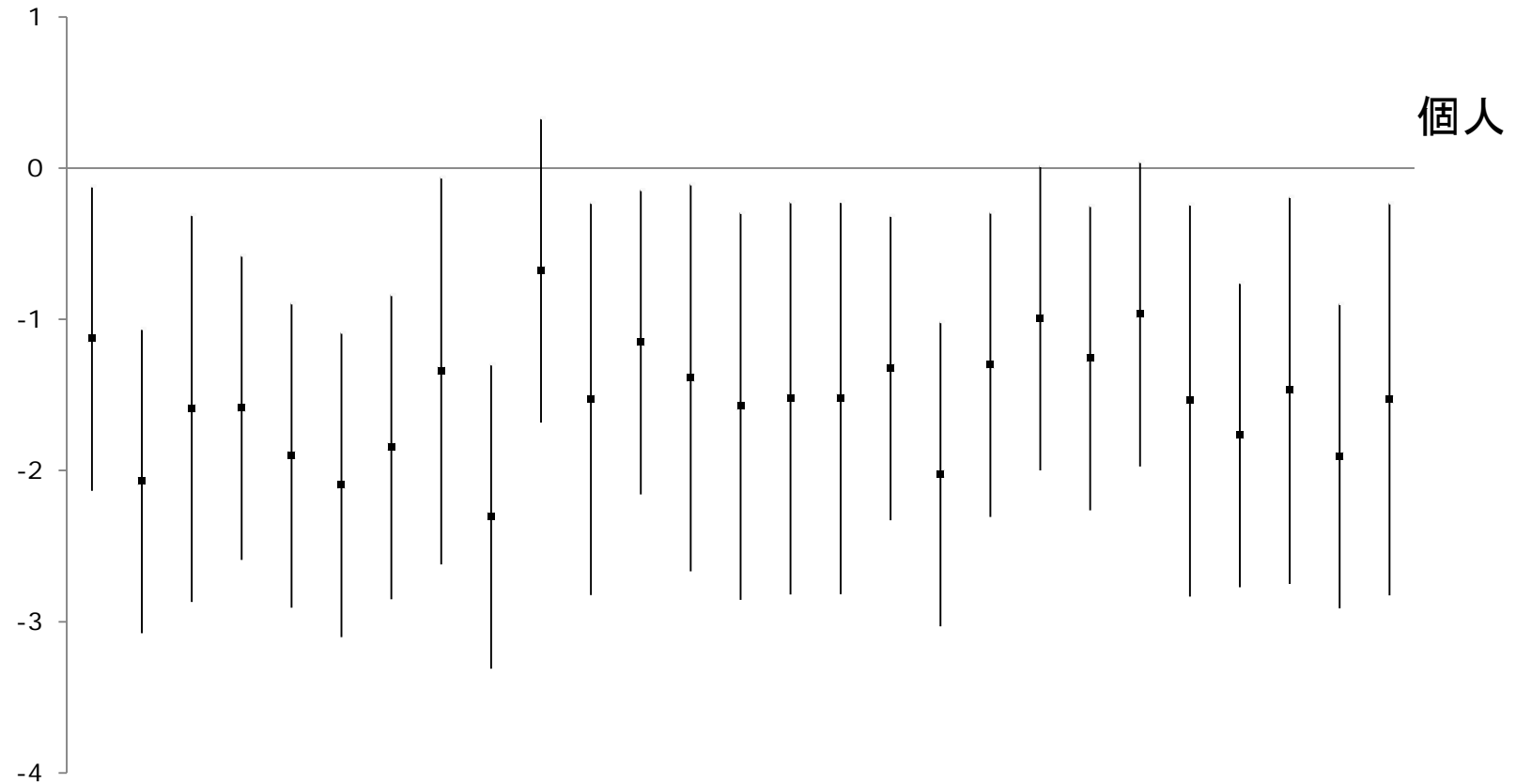
個人差の検討

- 切片に関する個人間のばらつき(個人差)
 - 分散 14.26 (SE 3.8)
- 傾きに関する個人間のばらつき(個人差)
 - 分散 0.39 (SE 0.6)
- 切片と傾きの関係
 - 共分散 0.096 → 相関係数 0.041
- 各個人の切片と傾きの予測
 - 反応しやすい(しにくい)個人の同定

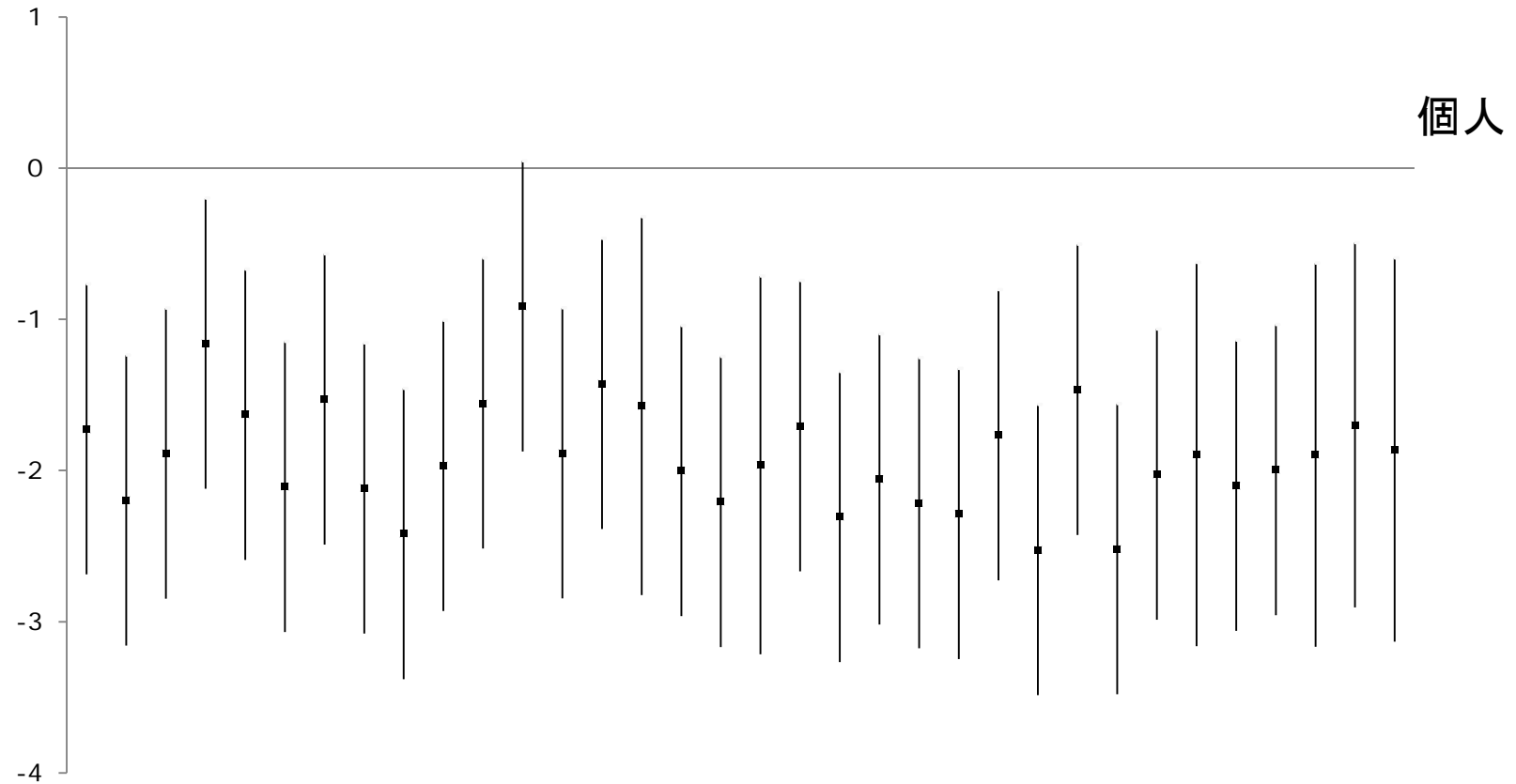
切片の個人差(予測値) 集団平均からのズレ



傾きの予測値 プラセボ群



傾きの予測値 エストロゲン群



変量効果の選択

- 混合効果モデル
 - 個人(クラスター)間のバラツキが大きいデータ
- 変量効果
 - 個体に特有の効果
 - 個体差を生じるレベルごと(理由は同定できなくてもよい)
- 経時データ
 - 切片に対する変量効果
 - 時点ゼロ(ベースライン)での個人差
 - 傾きに対する変量効果
 - 治療、あるいは変化に対する個人差
- 形式的な検討よりも、臨床的・疫学的に解釈可能な変量効果を取り込む

結果変数間の相関

- 集団平均アプローチ Population-Averaged
 - 個人内の結果変数間の相関: (作業)相関行列
 - 平均パラメータ(治療効果)の推定、検定に主眼
 - 相関構造には興味がない
- 対象特定アプローチ Subject-Specific
 - 個人内の結果変数間の相関: 変量効果
 - 回帰パラメータの異質性
 - 個人差、施設間差等の原因の探索

周辺モデル

平均構造のモデル化とロバスト分散

- 結果変数の周辺分布のみをモデル化
- 周辺モデルの重要性(治療効果の推定が目的)
- 一般化推定方程式 (GEE) によるパラメータ推定
- 平均構造に対するモデルが正しければ、
仮定した相関構造が真の相関構造でなくても
 - 平均パラメータ(治療効果など)の推定はバイアスなし
 - ロバスト分散を用いれば治療効果の検定、推定は問題なし

混合効果モデル 変量効果の導入

- 個人内の結果変数間の相関は、個人特有の何らかの要因（変量効果）によって生じている
- 変量効果を与えたもとでは、個人内の結果変数は独立
- 変量効果に対して分布を仮定
- 意味のある変量効果

内容

- 経時データとは
- いくつかの例
- 単純な解析
- 2つの解析アプローチ
 - 周辺モデル
 - 混合効果モデル
- 欠測データ(ちょっとだけ)
- まとめ

欠測データ

- 本来は測定されるべきデータであったが、何らかの理由で測定されなかったデータ
- 通常のデータには欠測データがある
- 特に、経時データではつきもの

結果変数の欠測

- 結果変数の測定が不十分
 - 追跡不能、悪化・改善による中止、副作用による中止
 - 薬剤を継続して使用することができない
 - 治療効果は無効か悪化
 - 有効性の評価から除くことはできない
 - 除くと治療効果の比較にバイアス
 - ITTの原理:とにかく主要エンドポイントは測定

参考文献・資料

- Diggle PJ, Heagerty P, Liang KY, and Zeger SL. Analysis of Longitudinal Data. 2nd edition. Oxford Science Publication. 2002.
- Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. 2nd edition. Wiley Series in Probability and Statistics. 2011.
- Laird NM and Ware JH. Random-effects models for longitudinal data. Biometrics, 38: 963-974, 1982.
- Ware JH. Linear models for the analysis of longitudinal studies. The American Statistician 39: 95-101, 1985.
- Zeger SL and Liang KY. An overview of methods for analysis of longitudinal data. Statistics in Medicine 11: 1825-1839, 1992.
- Dunlop DD. Regression for Longitudinal Data: A Bridge from Least Squares Regression. The American Statistician 48: 299-303, 1994.
- Albert PS. Longitudinal data Analysis (repeated measures) in clinical trials. Statistics in Medicine 18: 1707-1732, 1999.
- 松山裕、林邦彦、佐藤俊哉、大橋靖雄. Generalized Estimating Equationsの理論と応用. 薬理と治療. 24: 2531-2542, 1996.
- 松山 裕、山口拓洋 編訳. 医学統計のための線型混合効果モデル: SASによるアプローチ. サイエンティスト社. 2001 .
- 講義資料(医学データ解析入門: 東北大学、Longitudinal data analysis: Universiteit Hasselt & KU Leuven)