欠測データの統計科学：基礎理論と実践的な方法論

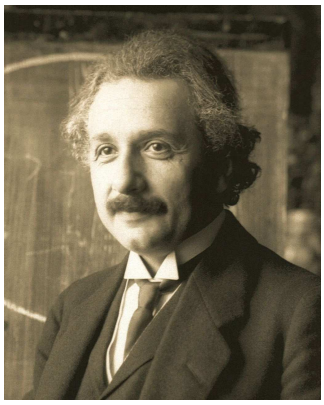
野間 久史
情報・システム研究機構　統計数理研究所
2017年1月13日
統計数理研究所 公開講座
e-mail: noma@ism.ac.jp
URL: http://normanh.skr.jp

1

手段の完璧さと、目的の混乱。この2つが、私たちにとっての主要な問題であるように思われる。

A perfection of means, and confusion of aims, seems to be our main problem.

—Albert Einstein



https://en.wikipedia.org/wiki/Albert_Einstein

2

Contexts

- 欠測データの統計解析：医学研究における応用から
  - National Research Councilによる調査報告書の推奨事項などを中心に、欠測データの統計解析の実践的な方法論の基礎について、（午前中の復習も兼ねて）解説を行います
- Case Study 1: Multiple Imputation by Chained Equation (MICE)
  - 複数の変数に非単調に起こった欠測を扱うには？？
- Case Study 2: 脱落を伴う経時測定データの統計解析
  - 非ランダムなメカニズムで起こる脱落によるデータの欠測をどのように扱えばよいのか？？

3

Missing Data

- 一般的な統計学の教科書で解説される統計手法は、原則としてすべて「欠測はひとつもなく、完全なデータが観測されている」ことを前提としている
- しかしながら、ほとんどすべての調査・実験研究において、なんらかのデータの欠測は生じる
  - 長期間の追跡調査における脱落，追跡不能（Drop-out）
  - 質問紙調査における無回答，無記入
  - 計測機器の測定限界を超えるデータ

など

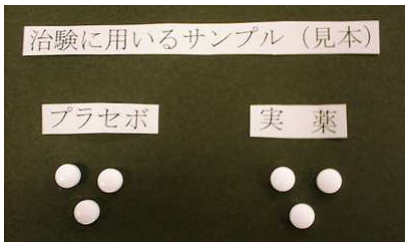
4

Lurasidone 第2相試験

- Lurasidone（商品名：Latuda）
- 統合失調症の治療薬
- 米国 Food and Drug Administration (FDA) において、2010年10月承認
- 第2相試験（D1050006）
  - プラセボ（n=49），40mg投与群（n=49），120mg投与群（n=47）の3群比較のランダム化比較試験

5

補足：プラセボとは？



http://medical.radionikkei.jp/medical/suzuken/final/021024html/index.html

見た目，味，匂いなどでは
まったく見分けがつかない

- ヒトの体には、とても不思議な一面があります。乳糖やでんぷんなど、くすりとしての効き目のないもので錠剤やカプセル剤をつくり、頭痛の患者に本物のくすりとして服用してもらう実験をすると、半数くらいの人が治ってしまうこともあります。くすり（に似たもの）を飲んだという安心感が、体にひそむ自然治癒力を引き出すのかもしれません。これを「プラセボ効果」といいます。プラセボは、一般に偽薬（ぎやく）と訳されていますが、くすりに似せた気安めのものといってもいいでしょう（http://www.takeda.co.jp/ct/placebo.html）
- 医薬品開発の臨床試験では、プラセボ効果によるバイアスを考慮した上で、試験薬の治療効果を正しく評価するために、プラセボ（偽薬）を比較対照として用いることが多い

6

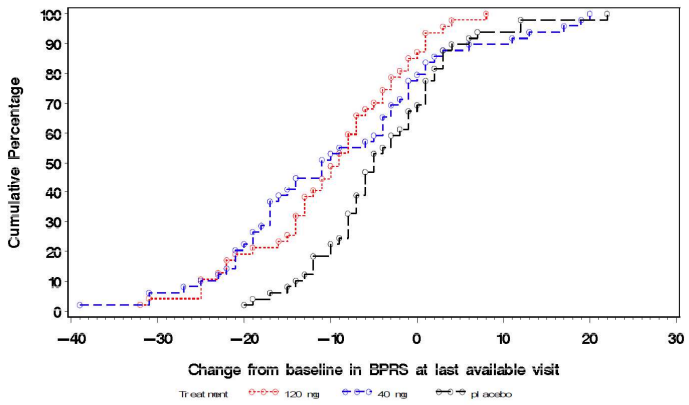
補足：ランダム化比較試験

- 字義通り、試験薬とプラセボを、対象者集団にランダムに割り付ける（コンピュータで生成した乱数などから）という試験デザイン
- もともとは、R. A. Fisherによって、農事試験における、交絡（見せかけの処理効果）を除いて、バイアスのない評価を行うために考えられた
- 仮に、医師や製薬企業が意図的に「状態の悪い患者に多くプラセボを割り付けた」としたら、結果として認められる「試験薬とプラセボの治療成績の差」は、「試験薬の有効性」だけではなく、単純に「試験薬群に状態の良い患者が多く、プラセボ群に状態の悪い患者が多かった」という対象者集団の偏りによるものかもしれない（＝見せかけの治療効果）
- このようなバイアスを防ぎ、正しい治療効果の評価を行うために、ランダム化比較試験という研究デザインが用いられる

7

Lurasidone 第2相試験の結果

Figure 1. Study D1050006 Empirical cumulative distribution functions



Source: Reviewer's results

日本製薬工業協会 (2014)

8

脱落（Drop-out）による欠測

表 2　BPRS Total Score LS Mean Change from Baseline（MMRM analysis）
（審査官による解析．Statistical Review（s），Table 8）

	Placebo		SM-13496 40 mg		SM-13496 120 mg	
Day	No	LS Mean (SE)	No	LS Mean (SE)	No	LS Mean (SE)
3	44	-3.2 (1.03)	47	-4.6 (1.01)	44	-5.0 (1.01)
7	44	-4.7 (1.37)	47	-6.2 (1.34)	44	-6.9 (1.37)
14	34	-5.5 (1.50)	32	-11.0 (1.53)	37	-11.5 (1.44)
21	26	-5.2 (2.00)	26	-13.2 (2.00)	24	-10.0 (2.01)
28	18	-6.5 (1.93)	22	-12.5 (1.87)	23	-12.8 (1.83)
35	15	-5.9 (1.92)	17	-13.2 (1.88)	21	-12.7 (1.78)
42	17	-4.1 (2.11)	17	-13.4 (2.10)	19	-13.4 (2.00)
Day 42	LS Mean Difference (SE)		-9.3 (2.95)		-9.2 (2.89)	
	Unadjusted p-value		0.0025		0.0022	

42日目までの脱落は70%近く発生！

日本製薬工業協会 (2014)

9

脱落を起こす前のアウトカム

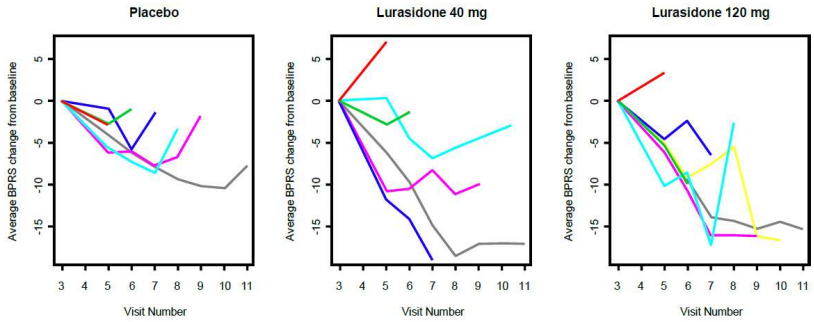


図 1　BPRS total score response profiles by treatment group　日本製薬工業協会 (2014)
（審査官による解析．Statistical Review（s），Figure 2 より改変）

脱落前にスコアが高くなる症例がほとんど！！

10

欠測による問題①：バイアス

- 欠測が何の理由もなく「完全にランダムに」起こってくれるのであれば (MCAR) 、欠測を無視して解析を行っても、治療効果の妥当な推定・検定を行うことができる
- しかしながら、一般的な医学研究で欠測を起こす対象者は「欠測を起こすなんらかの理由」がある
- 例えば、「試験治療を受けていたが、症状が悪化したため」「有害事象が起こったため」に脱落を起こすのであれば、これはランダムな脱落ではない
- 仮に、群間でのアウトカムの差が認められたとしても、それが群間の欠測パターンの違いによるものであったら…？？

11

欠測による問題②：推定精度・検出力の低下

- 欠測が起きると、統計的な推定・検定を行う上での情報量の損失が生じる
- せっかく多大な費用・労力をかけて行った研究でも、最終的な評価においては、欠測したデータの情報量の損失の分だけ、推定精度・検出力が低下してしまう
- 意味のある差があっても、欠測が多くなってしまうと、有意差が示せなくなってしまう可能性も

12

欠測による問題③：統計理論の想定外？？

- 共分散分析（線形回帰分析）

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

- 最小二乗法の二乗誤差関数

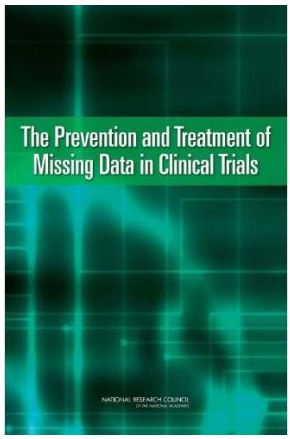
$$LS(\boldsymbol{\beta}) = \sum_{i=1}^{n} \{Y_i - (\beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2})\}^2$$

- 結果変数，説明変数の組のうち、1つでも欠測があれば、そもそも通常の最小二乗法（最尤法）における二乗誤差（尤度）が定義できない…？？

13

FDAの欠測データガイドライン

- 米国の規制当局（FDA）が医薬品開発における臨床試験での「欠測データの取り扱い」に関するガイドラインを作成することになった
- 米国 National Research Council (NRC) が、当該研究領域のエキスパートによる調査委員会を組織し（委員長：Roderick Little教授）、ガイドライン作成のための学術調査報告が行われた
- 解析手法だけではなく、欠測は、深刻なバイアスの原因となるため、あらかじめ最小限に抑えるための防止策と適切に取り扱うための包括的なRationaleが重要！！



http://www.nap.edu/catalog.php?record_id=12955

14

NRC Report: Recommendations

- 1. 試験のプロトコルにおいては、以下の4つの点が明確に定められるべきである。(a) 試験の目的，(b) 主要なアウトカム，および，副次的なアウトカム，(c) いつ、誰の、どのようなアウトカムを測定するか、(d) 主たる関心のある，推定の対象（estimands）となる介入効果の指標。これらの事項は、すべての対象者について意味のあるべきもので、最小の理論的仮定によって推定できるものであるべきである。後述のことにも関連して、プロトコルには、それらの仮定の潜在的な影響と欠測データの取り扱いについて明記されるべきである。
- 6. 試験のスポンサーは、欠測データが生じることによる潜在的な問題を明確に予想しておく必要がある。特に、プロトコルには、どの程度の欠測データが発生するかの見積もりや、欠測データの影響を最小限に留めるための試験計画、実施段階でのモニタリングの手順について明記した章も設けるべきである。

National Research Council (2010) 15

NRC Report: Recommendations

- 8.すべての試験のプロトコルは、欠測データを最小化することの重要性についての理解のもと、作成されるべきである。特に、過去の同様の試験の情報をもとにして、主要なアウトカムの情報の収集における最低限の達成目標を明確に定めるべきである。
- 9.試験のスポンサーは、欠測データの取り扱いに関する統計解析の方法について、プロトコルに明確に定めるべきである。また、それに関連する理論的な仮定についても、統計家だけではなく、臨床家が理解することができるように説明がされるべきである。
- 17. 米国FDAと、臨床試験のスポンサーとなる製薬企業，医療機器・生物学的製剤の製造販売企業は、所属する統計解析担当者が、欠測データの解析における最新の方法論に関する知見をフォローできるように、継続的にトレーニングを行うべきである。FDAは同様に臨床担当の審査官に対して、欠測データに関する専門用語や方法論に広く馴染めるように、継続的なトレーニングを行うことを推奨するべきである。

National Research Council (2010) 16

医学ジャーナルにおける動向

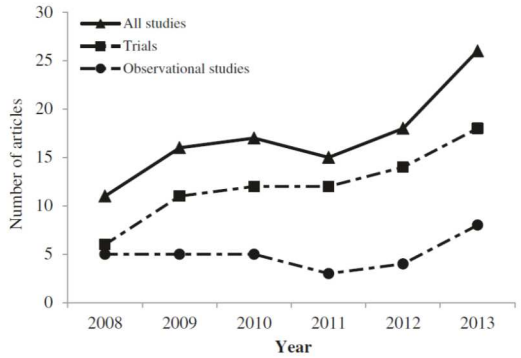


Figure 2 Number of articles in the Lancet and New England Journal of Medicine that used MI: overall and by study type.

- 多重代入法（multiple imputation; MI）に限った統計ではあるが、NRCによるRecommendationsが公表されて以降、一流医学ジャーナルの中でも、Complete-Case AnalysisのようなRoughな方法から、一定の理論的仮定のもとでの妥当な方法論へと、スタンダードとなる方法がシフトしつつある

Rezven et al. (2015)

17

NRC ReportによるRecommendationsから

- 一般的に、欠測データの統計解析は高度な方法論・計算技法を用いたものが多く、また統計ソフトウェアにも標準的なモジュールとして実装されていないものが多かった
- 実質的には、学術誌の査読でも、不適切な解析（Complete-Case Analysisなど）が行われていても、見逃されてきた
- しかしながら、NRC ReportによるRecommendationsは、データサイエンスのコミュニティにおいて、この問題を再考する大きな契機となり、この数年ほどで、さまざまな領域で方法論に関する議論が行われるようになり、ソフトウェアの開発なども大きく進められた

18

欠測データの統計解析：Rationale

- まず、臨床試験において、すべての欠測データを統一的に扱う方法は存在しない
- 個々の試験のデザイン，測定値の特性などに応じて，必要な仮定・モデルは違う
- モデリングや推測の方法も広範に及び、いかなる状況においても万能な方法は存在しない
- Case-by-caseで、適切な解析手法を的確に選択する必要がある（それぞれの手法に対して正確な知識が必要！！）

Little et al. (2012) 19

欠測データの解析における4つのアプローチ

- Complete-Case Analysis
- 単純な代入方法（Single Imputation）
- 重みつき推定方程式（Weighted Estimating Equation; WEE）による方法
- モデルに基づく方法（最尤法，ベイズ推測，多重代入法など）

Little et al. (2012) 20

Complete-Case Analysis

- 主要なモデルにおいて、少なくとも1つの変数が欠測している対象者を、単純に、解析対象集団から除外する
- 残された対象者は、すべての変数が測定されているので、形式的に、完全データのデータセットができ上がる
- 通常の完全データに対する解析手法（最小二乗法，最尤法など）を適用することができる

21

単純な代入方法

- 欠測値に対して、適当な単一の値を代入する方法（Single Imputation）
- Last Observation Carried Forward (LOCF)
  - 脱落を起こした時点での値を、単純に代入値として利用する単一代入法
- 日本で行われる治験でも、LOCFは慣習的に多くの試験で用いられてきた

22

LOCF (Last Observation Carried Forward)

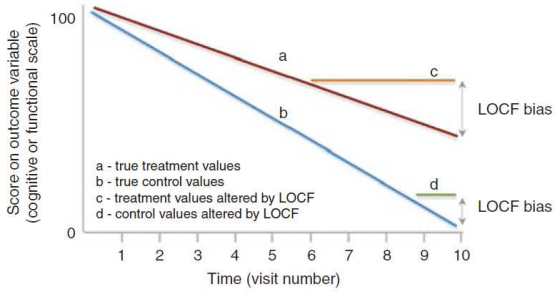


Figure 2 Differential last-observation-carried-forward (LOCF) bias when there are more or earlier dropouts in the treatment group than in the control group (effect measures by LOCF [c–d] > true effect [a–b], resulting in an exaggerated positive effect, biased in favor of treatment). Figure extracted from ref. 4.

O'Neill and Temple (2012)

最終観測時点での測定値から、アウトカムが不変であると仮定。

23

WEE (Weighted Estimating Equation)

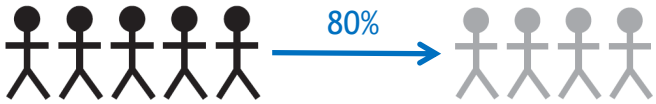
- 完全データが観測された対象者に対しての重みつき推定方程式
- 「観測される確率（欠測を起こさない確率）の逆数」で重みつけた推定関数に基づく推定量
- いわゆる IPW (Inverse Probability Weighting) 推定量
  - 推定関数の不偏性から、推定量の一致性は保証される

24

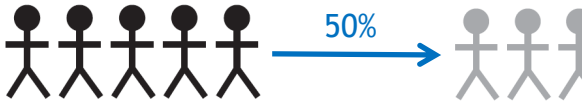
個々人が持つデータの情報量は？？

全集団　　観測データ

欠測を起こしにくい人



欠測を起こしやすい人



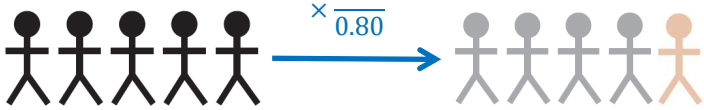
推定量への貢献度を調整すれば？

25

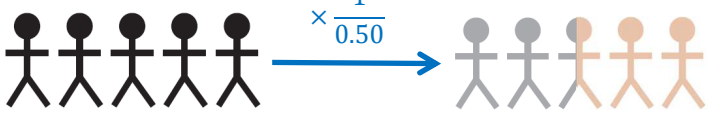
個々人が持つデータの情報量は？？

全集団　　観測データ

欠測を起こしにくい人



欠測を起こしやすい人



同じ情報量を持つ
データを推定関数上で
'Impute' するという
考え方！！

個々人の重み（寄与率）を調整することで
欠測メカニズムによって生じるバイアスを補正

26

IPW法

- 「真の観測確率」に基づく重みよりも、全員に観測されているベースライン共変量などを利用して、2項回帰モデルなどで「推定した観測確率」を重みに用いたほうが推定精度は必ず高くなる
- 欠測を起こした対象者の情報も組み込んで、検出力をUPできる!!
  - cf.) より効率の高い方法として、Augmented IPW法（Doubly Robust推定法; セミパラメトリック有効）という方法もある

27

モデルに基づく方法①

- 最尤法（Maximum Likelihood; ML）
- 欠測を起こした変数の分布にパラメトリックな確率分布モデルを仮定する
- 欠測変数の分布に対して、尤度関数を積分
- 観測データに関する周辺尤度が得られる
  - Directに周辺尤度を最大化（Direct ML）
  - EMアルゴリズム

Dempster et al. (1977), Little and Rubin (2002)

28

モデルに基づく方法②

- ベイズ流の方法（Bayesian Methods）
- 欠測を起こした変数の分布にパラメトリックな確率分布モデルを仮定する（＋事前分布）
- MCMCを使えば、関心のあるパラメータの周辺事後分布を簡単に求められる
- 欠測変数の分布の仮定が正しければ、Direct MLと同様、妥当な事後推測が可能に
  - Data Augmentation法など

Schafer (1997)

29

モデルに基づく方法③

- 多重代入法（Multiple Imputation; MI）
- 欠測値に対して複数の代入値（M組）を生成
  - 代入値の生成方法はいろいろ（後ほど）
- M組の擬似的な完全データに対して、推定値と分散の推定値を求め、Rubinによる統合公式によって統合
  - $\hat{\beta}_h, \hat{V}(\hat{\beta}_h), h = 1,2,\ldots,M$

Rubin (1987)

30

Rubinの統合公式

- M 回の推定結果を統合

$$\hat{\beta}_{\mathrm{MI}} = \frac{1}{M}\sum_{h=1}^{M}\hat{\beta}_h$$

$$\hat{\mathrm{V}}(\hat{\beta}_{\mathrm{MI}}) = \underbrace{\frac{1}{M}\sum_{h=1}^{M}\hat{\mathrm{V}}(\hat{\beta}_h)}_{\text{完全データの推定量の分散}} + \underbrace{(1+M^{-1})\frac{\sum_{h=1}^{M}(\hat{\beta}_h-\hat{\beta}_{\mathrm{MI}})(\hat{\beta}_h-\hat{\beta}_{\mathrm{MI}})^T}{M-1}}_{\text{欠測値の予測の不確実性によって生じる付加的なばらつきを表す項}}$$

Rubin (1987)

31

Revisited：単一代入法

- すべての欠測値に対して、100%の確率で「本当は観測されるはずだった値」を予測できれば、完全データを再現できる
- しかし、100%で予測できることはまずありえない（できるなら、その値に置き換えてやればよい）ので「代入値の予測の不確実性」は必ず生じる
- 代入値の不確実性は、最終解析の推定値の不確実性（分散）にも影響するはず
- 完全データに対するモデルの分散の推定量は、この不確実性を反映したものとなっていない

32

Revisited：単一代入法

- 多重代入法の分散公式

$$\hat{\mathrm{V}}(\hat{\beta}_{\mathrm{MI}}) = \underbrace{\frac{1}{M}\sum_{h=1}^{M}\hat{\mathrm{V}}(\hat{\beta}_h)}_{\text{完全データの推定量の分散}} + \underbrace{(1+M^{-1})\frac{\sum_{h=1}^{M}(\hat{\beta}_h-\hat{\beta}_{\mathrm{MI}})(\hat{\beta}_h-\hat{\beta}_{\mathrm{MI}})^T}{M-1}}_{\text{欠測値の予測の不確実性によって生じる付加的なばらつきを表す項}}$$

- 代入後の分散の推定量を、通常の完全データのモデルの分散の推定量とすると、上式の第1項しか考慮しないものとなる（過小推定のバイアスが生じる）
- 対応する検定のP値，信頼区間も誤り

33

単一代入法と多重代入法

- 多重代入法で生成する代入値のデータとは？
  - 欠測したデータの背景にある確率分布（真の構造）から生成されるデータ（いわゆる乱数）
- 1点としての「正確な値」は予測できないが、観測されたデータから、欠測したデータが「どのような分布に従って得られるはずのものであったか」を推定することはできる
- その分布からの値を、複数シミュレーションして、その「分布の情報」を補完値として組み込む（というイメージ）

34

欠測データに対する3つのシナリオ

- MCAR (Missing Completely At Random)
- MAR (Missing At Random)
- MNAR (Missing Not At Random)

Little et al. (2012)

35

MCAR

- すべての欠測は、完全にランダムに起こる（いかなる変数ともまったく無関係）
- Complete-Case Analysisで妥当な結論が得られる
  - 解析対象集団から、ランダムに一定の割合の対象者を除外することと同じ（除外後のデータも、母集団からの偏りのないランダムサンプルと見なすことができる）
  - ただし、検出力の低下は起こる

36

Complete Case Analysis

- 「MCARは、極めてあり得ない仮定である」
- 臨床試験で、脱落や追跡不能が起こる場合、「何の理由もなくランダムに」という都合のよい仮定はまずあり得ない
- 脱落を起こす患者は、一般的に、脱落を起こすなんらかの理由がある（症状の悪化，副作用など）
- NRCレポートでは「Complete-Case Analysisは推奨しない」と明言されている

Little et al. (2012)

37

Lurasidoneの第2相試験

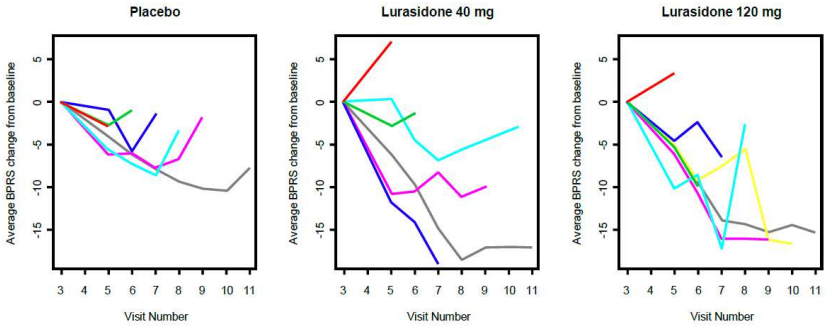


図 1　BPRS total score response profiles by treatment group

（審査官による解析．Statistical Review（s），Figure 2 より改変）

日本製薬工業協会 (2014)

脱落前に症状が悪化している!!

38

MAR

- 欠測のメカニズムは、観測されている変数ですべて完全に説明することができる
- WEEやモデルに基づく推定方法（ML, Bayes, MIなど）で、妥当な推測が可能
- 単一代入法（LOCFなど）も仮定が正しければ妥当な（もしくは保守的な）評価が可能??
  - 単一代入法を利用する上では、十分な科学的根拠を提示する必要がある、とされている

39

Example. LOCF

- 脱落を起こした対象者のアウトカムは、最後まで脱落時点から不変である
  - この仮定を支持する、科学的根拠は提示できるか？
  - 提示できなければ、LOCFによる科学的妥当性は担保されない（非常に強い仮定に依存した方法であるため）
- LOCFのような解析手法を、主要な解析に利用する場合は、その科学的根拠を明確に説明できなくてはいけない

40

単一代入法

- すべての単一代入法は「欠測データを100%確実に予測できる」のでなければ、分散を過小推定する
- 得られるP値も誤り（Type-1 Error Rateを名目水準以下に保持できない）
- 得られる95%信頼区間も誤り（狭すぎる；真値を含む確率は95%にはならない）
- 主要な解析に用いるのであれば、この精度の問題も含め、科学的な根拠を説明できなくてはいけない

Little et al. (2012)

41

MNAR

- 欠測のメカニズムは、観測されている変数では完全に説明することができない
- 観測されていない変数にも影響される
- 感度解析をするしかない!!
- Pattern-Mixture Models
- Selection Models
- MNARのもとでの解析手法は、やや高度な話題であり、このセクションでは詳細については解説しません
- ご関心がおありの方は、高井ら (2016)，日本製薬工業協会 (2016) などをご参照ください

42

Favored Methods

- WEEとモデルに基づく方法を推奨！！
- 欠測を起こした対象者において、最終解析に組み込めない補助的な情報（ベースライン共変量など）を解析に取り込み、推定精度の改善などに役立てることができる
- 欠測データの不確実性を反映したP値や信頼区間を与えることができる（単一代入法とは異なる）

Little et al. (2012) 43

Sensitivity Analysis

- MARを仮定した方法にも、欠測メカニズムや欠測変数の分布を規定する付加的な仮定を置く必要がある
- これらは、観測されたデータからはけっして検証することができない仮定である
- 感度解析を行うことが重要（パターン混合モデルなど）

Little et al. (2012) 44

NRC Report: Recommendations

- 10. Last observation carried forward (LOCF) や Baseline observation carried forward (BOCF) などの単一補完法は、それらの背景にある理論的仮定が科学的に正当化されない場合には、主要な解析方法として用いられるべきではない。
- 14. 欠測データの発生が予想される場合、欠測メカニズムや関心のあるアウトカムに関連すると思われる補助的な情報を収集するべきである。これらの情報は、MARの仮定のもとで、解析モデルの構築に利用することができ、推測の妥当性と精度の改良のために有用な情報となる。また、感度解析においても、治療効果の推定における欠測データの影響を評価するために利用することができる。加えて、脱落を起こした参加者の全員、もしくは、ランダムに選択された一部の者に対して、なぜ脱落を起こしたかの理由を調べることも検討されるべきである。また、もし承諾が得られるのであれば、その脱落を起こした後のアウトカムの情報も収集するべきである。
- 15. 感度解析は、臨床試験の主要な報告の一部として位置づけられるべきである。欠測データのメカニズムに関する理論的仮定の感度を評価することは、必須の報告事項とされるべきである。

45

Lurasidone 第2相試験

- 試験治療の中止割合が70%近くもあり、他の統合失調症の試験に比べても非常に高く、審査官の間でも議論を呼んだ
- 事前の解析計画では、LOCF-ANCOVA (Analysis of Covariance) での解析が予定されていたが、42日目の測定値が得られている対象者はわずか3割であり、それ以外の対象者は、全員、脱落時点での測定値を強制的に代入することに

46

主要な解析の結果

表 1　BPRS Total Score LS Mean Change from Baseline to Week 6（LOCF-ANCOVA）
（申請者による解析．Statistical Review（s），Table 6）

		Placebo	**SM-13496 40 mg**	**SM-13496 120 mg**
No patients	N=145	49	49	47
Baseline	Mean (SD)	54.7 (8.13)	54.2 (8.93)	52.7 (7.61)
Change from Baseline	LS Mean (SE)	-3.8 (1.57)	-9.4 (1.58)	-11 (1.58)
Placebo-adjusted Difference	LS Mean Difference	NA	-5.6 (2.13)	-6.7 (2.16)
	95% CI	NA	(-9.8, -1.4)	(-11, -2.5)
	p-value (Dunnett)	NA	0.018	0.004

日本製薬工業協会 (2014)

47

FDA審査官による感度解析

表 2　BPRS Total Score LS Mean Change from Baseline（MMRM analysis）
（審査官による解析．Statistical Review（s），Table 8）

	Placebo		**SM-13496 40 mg**		**SM-13496 120 mg**	
Day	**No**	**LS Mean (SE)**	**No**	**LS Mean (SE)**	**No**	**LS Mean (SE)**
3	44	-3.2 (1.03)	47	-4.6 (1.01)	44	-5.0 (1.01)
7	44	-4.7 (1.37)	47	-6.2 (1.34)	44	-6.9 (1.37)
14	34	-5.5 (1.50)	32	-11.0 (1.53)	37	-11.5 (1.44)
21	26	-5.2 (2.00)	26	-13.2 (2.00)	24	-10.0 (2.01)
28	18	-6.5 (1.93)	22	-12.5 (1.87)	23	-12.8 (1.83)
35	15	-5.9 (1.92)	17	-13.2 (1.88)	21	-12.7 (1.78)
42	17	-4.1 (2.11)	17	-13.4 (2.10)	19	-13.4 (2.00)
Day 42	LS Mean Difference (SE)		-9.3 (2.95)		-9.2 (2.89)	
	Unadjusted p-value		0.0025		0.0022	

日本製薬工業協会 (2014)

48

審査官のコメント

- 審査官のコメント：他の統合失調症の試験と比べて，本試験では中止割合が約70%と非常に高いが，このことにより試験結果が解釈不能になるわけではないと考える．中止の内容は，少なくとも予想された傾向（全中止はプラセボ投与群で最も多く，効果不十分による中止もプラセボ投与群で最も多い）であった．加えて，MMRMとObserved Case（OC）の解析結果はLOCFの結果を（少なくとも数値としては）支持するものであり，Day3以降のいずれの時点においても本剤ではプラセボよりも改善がみられた．副次評価項目（CGI-S）の結果でも，本剤ではプラセボよりも大きく改善した．以上より，中止割合が高いことは問題ではあるものの，本試験はSM-13496（Lurasidone）について肯定的な試験と考える．

日本製薬工業協会 (2014)　49

Contexts

- 欠測データの統計解析：医学研究における応用から
  - National Research Councilによる調査報告書の推奨事項などを中心に、欠測データの統計解析の実践的な方法論の基礎について、午前中の復習も兼ねて解説を行います
- Case Study 1: Multiple Imputation by Chained Equation (MICE)
  - 複数の変数に非単調に起こった欠測を扱うには？？
- Case Study 2: 脱落を伴う経時測定データの統計解析
  - 非ランダムなメカニズムで起こる脱落によるデータの欠測をどのように扱えばよいのか？？

50

卵巣がんの予後因子研究

- がんの臨床研究では、患者の予後（症状の悪化や生存期間など）と関連する因子を同定することは、 症状の経過や治療の選択に重要な情報となるため、ひとつの大きな研究テーマとなっている
- Clark et al. (2001) は、英国 Edinburghの Western General Hospital の診療データベースをもとに、1189人の卵巣がん患者のデータ（診断年：1984-1999年）から、予後因子の同定および予後予測モデルの構築に関する研究を行っている

51

データベースの2次利用研究

- データベースを2次利用した「後ろ向き（retrospective）」に行われる研究では…
- あらかじめ、どのような研究を行い、統計解析にどのような変数の情報が必要になるか、明確に定められたもとでデータが集められたわけではないので、重要な変数にしばしば大きな割合で欠測が含まれる
- Clark et al. (2001) の予後因子研究でも、多くの変数に多数の欠測が含まれることが問題となっていた

52

欠測データの内訳（n=1189）

Variable	Name	Type	Levels	% missing
Albumin (*outcome variable*)	`alb`	Continuous (rounded)	30	33.0
Grade of tumour	`grade`	Ordinal	3	11.6
Residual disease	`resdis`	Ordinal	3	6.8
Performance status	`ps`	Ordinal	4	42.7
Presence/absence of ascites	`ascites`	Binary	2	5.4
Age (exact years)	`age`	Continuous	–	0.0
FIGO stage	`figo`	Nominal	4	1.8
Histology	`histol`	Nominal	7	0.0
Chemotherapy regimen	`ctype`	Nominal	3	0.0
Surgery (yes/no)	`surg`	Binary	2	0.0
CA125 (a cancer antigen)	`ca125`	Continuous	–	36.7
Alkaline phosphatase	`alp`	Continuous	–	33.1
All variables	–	–	–	70.1

Table 2: Variables and their missingness in the ovarian cancer dataset.

Royston and White (2011)

53

生存時間解析の回帰モデル

- Coxの比例ハザード回帰モデル（Cox, 1972）

$$\lambda(t) = \lambda_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p)$$

  - ハザード（hazard; 瞬間のイベント発生率）について、結果変数（死亡までの時間）と共変量間の関数関係を規定した回帰モデル
  - 時間に関するコンポネント $\lambda_0(t)$ に特定の関数形を仮定しなくても、$\beta_1,\ldots,\beta_p$についての推測を行うことができる、セミパラメトリックモデル（打ち切りがある場合にも、推定可能）
  - $\beta_1,\ldots,\beta_p$ は相対リスク（対数ハザード比）と解釈することができるパラメータ

54

Complete-Case Analysis

- 最も単純なアプローチとして、「結果変数と対象となる共変量のうち、少なくとも1つの変数が欠測している患者を除外した解析」を行うことを考えよう
- 238人（19.8%) の患者は、4つ以上の変数が欠測していた
- 少なくとも1つの変数が欠測している患者に至っては、831人（69.9%) にまで及び、Complete-Case Analysis を行う際のサンプルサイズは、1189人から358人（30.1%）にまで減少してしまう！！

55

MCARの妥当性…??

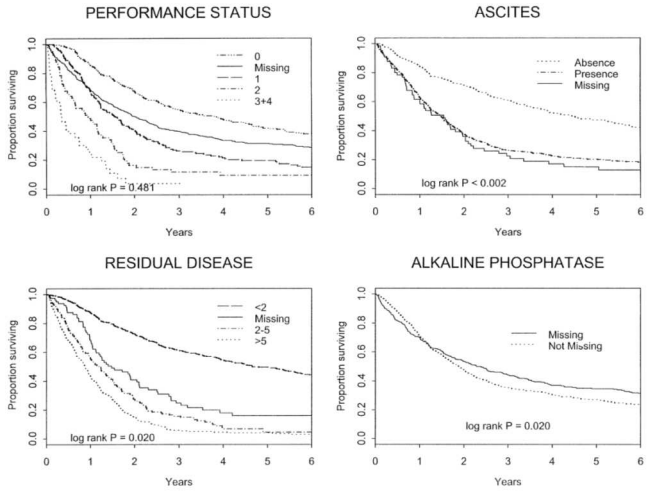


Fig. 1. Differences in survival of patients with and without recorded performance status, ascites, residual disease, and alkaline phosphatase, tested using log rank methods.

カテゴリごとのKaplan-Meier曲線と「欠測を起こした対象者」のKaplan-Meier曲線の比較

「完全にランダム」なメカニズムで欠測が生じているとは考えにくい…

56

補足：「欠測カテゴリ」を作るのはダメ？？

- 研究論文などを見ていると、カテゴリカル変数の欠測に対して、多重代入法などを行うのではなく、「欠測」というカテゴリを追加して解析をしているものを見かけますが、これは問題ないのでしょうか？
- 例えば、A, B, Cという3つのカテゴリがあり、実際には、1/3ずつそれぞれのカテゴリにデータが分布するものとします
- このとき、それぞれのカテゴリで、20%, 30%, 50%という頻度で欠測が起こったとしたら、どうでしょうか？単純に、「欠測カテゴリ」を設けて集計したとすると、真値である (1/3, 1/3, 1/3) からかなりかけ離れた結果が得られることは明らかです
- 基本的には、MCARのもとでなくては妥当な解析方法にならないため、推奨はできません！！

57

多重代入法（Multiple Imputation）

X_1	X_2	X_3	X_4	X_5
0	-0.125	1.129	0.049	1.084
1	0.694	0.602	1.018	NA
0	-0.761	1.229	0.922	-0.343
0	-0.809	-1.464	1.089	0.870
1	0.327	-1.527	-1.459	NA
1	-0.243	-1.488	-1.449	-1.132
…	…	…	…	…

-1.240
-0.898
⋮
-1.338

1.890
0.911
⋮
1.084

1つではなく、複数の代入値を用いる！！

58

代表的な代入値の生成方法

- 連続変数：線形回帰モデル，予測平均マッチング
- カテゴリカル変数（2水準）：ロジスティック回帰モデル
- 順序を持たないカテゴリカル変数（3水準以上）：多項ロジスティック回帰モデル
- 順序を持つカテゴリカル変数（3水準以上）：多項ロジスティック回帰モデル，比例オッズモデル
- その他：マルコフ連鎖モンテカルロ法，ベイズ流（近似）ブートストラップ法　など

White et al. (2011)

59

線形回帰モデルによる代入値生成

- 欠測した変数を結果変数として、観測されているデータで、その分布を説明する回帰モデルを作る

$$X_1 = \gamma_1 + \gamma_2 x_2 + \gamma_3 x_3 + \cdots + \gamma_q x_q$$

- 最小二乗法によって推定された回帰式に、推定された誤差分散を上乗せさせて、X_1 が欠測した患者の X_1 の予測値を乱数によって発生させる
- この予測値を、代入値として用いるという方法

60

予測平均マッチング

- 回帰モデルによる方法で推定された回帰式と誤差分散には誤差が伴う（誤差の分布が正規分布から外れることも）
  - Plausibleでない値が生成されることも当然ある
- よりPlausibleな値に近づけるために、推定された回帰式から予測される「予測平均」に近い観測値が得られている他の対象者（例えば、近い順に5人）のデータから、補完値をランダムに選びとる（マッチングさせる）という方法
  - *Hot-deck Imputationといわれる補完方法*
- シミュレーションなどによる経験的な評価では、相対的に良好な性能を持つ方法であることが知られている

61

ロジスティック回帰モデルによる代入値生成

- 線形回帰モデルと同じく、欠測した変数を結果変数として、観測されているデータで、その分布を説明する回帰モデルを作る

$$\text{logit}\ \Pr(X_1 = 1) = \gamma_1 + \gamma_2 x_2 + \gamma_3 x_3 + \cdots + \gamma_q x_q$$

- 最尤法によって推定された推定値 $\hat{\boldsymbol{\gamma}}$ とその漸近共分散行列の推定値 $\hat{\boldsymbol{V}}$ に対して、$MVN(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{V}})$ から乱数 $\boldsymbol{\gamma}^{*(1)}, \ldots, \boldsymbol{\gamma}^{*(M)}$を発生させ、これをplug-inした上記のロジスティック回帰の予測式から、欠測した X_1 の予測値（代入値）を生成する
- 多項ロジスティック回帰，比例オッズモデルも同様

62

単純な代入値の生成方法の問題点

- ここまで述べてきた、回帰モデルによる方法などでは、原則として、欠測を起こす変数が1つのみであることを想定している（SAS, Rなどに実装されているソフトウェアでも）
- しかし、一般的な調査・実験研究では、欠測を起こす変数が1つのみという性質のよい状況は、ほとんどあり得ない
- Clark et al. (2001) の卵巣がん研究でも、患者ごとにさまざまな組み合わせで複数の変数に欠測が起こっていた
- この場合、「欠測変数の予測のための回帰モデル」の説明変数が欠測することになり、欠測値の予測のためのモデルでも、また欠測に悩まされることになってしまう？？

63

複数の変数に欠測がある設定 (Example)

X_1	X_2	X_3	X_4	X_5
0	NA	1.129	0.049	1.084
1	0.694	NA	1.018	-1.240
NA	-0.761	1.229	0.922	-0.343
0	-0.809	-1.464	1.089	NA
NA	NA	-1.527	-1.459	1.084
1	-0.243	-1.488	NA	-1.132
…	…	…	…	…

64

連鎖方程式による多重代入法 (MICE)

- Multiple Imputation by Chained Equation (MICE)
- 複数の変数にまたがって、多重代入法による解析を行うための方法として開発された方法
- 基本的には、ここまで説明した多重代入法のアルゴリズムを組み合わせることによって実行することができる
- 個人ごとに、異なる組み合わせで複数の変数に欠測が起こっていても適用することができる
- 欠測を起こしている変数の型が異なっていて、異なる代入値の生成方法が必要とされる場合にも、適用することができる

White et al. (2011), Royston and White (2011)

65

MICEに必要となる仮定

- Fully Conditionally Specified (FCS) の仮定
  - $X_1, X_2, \ldots, X_p$という変数の組は、すべての変数が、他の $(p-1)$ 個の変数によって、お互いの分布を完全に説明することができるという仮定
- X_1の分布は、$X_2, X_3, \ldots, X_p$で完全に説明できる
- X_2の分布は、$X_1, X_3, \ldots, X_p$で完全に説明できる
- …

66

MICEのアルゴリズム①

- FCSの仮定のもとで、$X_1, X_2, \ldots, X_p$のそれぞれの変数についての補完値を順繰りに（連鎖的に）生成していく
- X_1についての代入値を生成する場合
  - $X_2, X_3, \ldots, X_p$ の欠測値には、1時点前の補完値を入れておき、その擬似的な完全データに対して、X_1の予測モデルを構築する
  - X_1 についての補完値を1組生成し（回帰モデルによる方法，予測平均マッチングなどで）、これをCurrentの値として更新する

67

虫食い状の不完全データ

X_1	X_2	X_3	X_4	X_5
0	NA	1.129	0.049	1.084
1	0.694	NA	1.018	-1.240
NA	-0.761	1.229	0.922	-0.343
0	-0.809	-1.464	1.089	NA
NA	NA	-1.527	-1.459	1.084
1	-0.243	-1.488	NA	-1.132
…	…	…	…	…

68

Initial Setting

X_1	X_2	X_3	X_4	X_5
0	0.694	1.129	0.049	1.084
1	0.694	-1.527	1.018	-1.240
1	-0.761	1.229	0.922	-0.343
0	-0.809	-1.464	1.089	-1.240
0	-0.809	-1.527	-1.459	1.084
1	-0.243	-1.488	0.049	-1.132
…	…	…	…	…

69

Cycle for X_1

X_1	X_2	X_3	X_4	X_5
0	0.694	1.129	0.049	1.084
1	0.694	-1.527	1.018	-1.240
NA	-0.761	1.229	0.922	-0.343
0	-0.809	-1.464	1.089	-1.240
NA	-0.809	-1.527	-1.459	1.084
1	-0.243	-1.488	0.049	-1.132
	…	…	…	…

X_1の予測モデルの構築

70

Cycle for X_2

X_1	X_2	X_3	X_4	X_5
0	NA	1.129	0.049	1.084
1	0.694	-1.527	1.018	-1.240
1	-0.761	1.229	0.922	-0.343
0	-0.809	-1.464	1.089	-1.240
0	NA	-1.527	-1.459	1.084
1	-0.243	-1.488	0.049	-1.132
…	…	…	…	…

X_2の予測モデルの構築

71

Cycle for X_3

X_1	X_2	X_3	X_4	X_5
0	0.521	1.129	0.049	1.084
1	0.694	NA	1.018	-1.240
1	-0.761	1.229	0.922	-0.343
0	-0.809	-1.464	1.089	-1.240
0	0.023	-1.527	-1.459	1.084
1	-0.243	-1.488	0.049	-1.132
…	…	…	…	…

72

MICEのアルゴリズム②

- 以上の連鎖的なアルゴリズムを $X_1, X_2, \ldots, X_p$ に順繰りに繰り返していき、M組の代入値の組を作成
  - Multiple Imputation by Chained Equation (MICE)
- 経験的に良好な性能を持つ方法であることも知られてきており、標準的なソフトウェアにも実装されてきている
  - SAS PROC MI: fcs statement
  - R, S-PLUS library: mice
  - Stata module: ICE

73

R example: Complete-Case Analysis

```
ph1 <- coxph(Surv(t, d) ~ age + figo + grade + histol + ascites + ps + resdis + log_ca125 + log_alp, data=tgce)
summary(ph1)

Call:
coxph(formula = Surv(t, d) ~ age + figo + grade + histol + ascites +
    ps + resdis + log_ca125 + log_alp, data = tgce)

  n= 362, number of events= 248
   (827 observations deleted due to missingness)

                                    coef exp(coef)  se(coef)      z Pr(>|z|)
age                             0.017297  1.017447  0.006798  2.545 0.010943 *
figoII                          0.869917  2.386712  0.377383  2.305 0.021159 *
figoIII                         1.388945  4.010617  0.325320  4.269 1.96e-05 ***
figoIV                          1.762204  5.825262  0.372662  4.729 2.26e-06 ***
grademoderately differentiated  0.248357  1.281918  0.344870  0.720 0.471434
gradepoorly differentiated      0.118153  1.125416  0.325479  0.363 0.716596
histolAdenocarcinoma            0.428714  1.535282  0.724775  0.592 0.554176
histolEndometrioid             -0.039115  0.961640  0.315038 -0.124 0.901188
histolMesonephroid (clear cell) 0.535352  1.708049  0.274731  1.949 0.051338 .
histolMixed mesodermal          1.383109  3.987278  0.410101  3.373 0.000745 ***
histolMucinous                 -0.265081  0.767144  0.173456 -1.528 0.126455
histolUndifferentiated          0.356664  1.428555  0.443565  0.804 0.421349
ascitespresent                  0.298751  1.348174  0.164309  1.818 0.069029 .
ps1                             0.150572  1.162499  0.152834  0.985 0.324525
ps2                            -0.032588  0.967938  0.225161 -0.145 0.884924
ps3to4                          0.832012  2.297938  0.443744  1.875 0.060795 .
resdis2-5 cm                   -0.028585  0.971820  0.190122 -0.150 0.880488
resdis<2 cm                    -0.567408  0.566993  0.189456 -2.995 0.002745 **
log_ca125                       0.062361  1.064347  0.049242  1.266 0.205366
log_alp                         0.477335  1.611773  0.175360  2.722 0.006488 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

ほとんどすべての統計ソフトは、デフォルトで
欠測データを除外し、Complete-Case Analysisを行う

74

R Example: MICE

```
library(mice)    # mice というパッケージを読み込む（install.package(“mice”) でインストールできる）

predmt1 <- (1 - diag(1, ncol(tgce)))
predmt1[c(1,11,12),] <- predmt1[,c(1,11,12)] <- 0
      # 代入値の生成モデルにおいて、それぞれの変数の予測モデルの説明変数の指定（行列形式）

imp.tgce <- mice(tgce, m=200, predictorMatrix=predmt1, seed=34871)
      # 補完値の生成方法については、次頁（ここではデフォルトの方法をそのまま採用）
      # m は、代入値の組の数（デフォルトでは5であるが、ここでは200としている）
      # predictorMatrixで、予測モデルに含める変数の組を指定している
      # seedは、乱数のシード

complete(imp.tgce, 5)   # 5番目のimputed datasetを出力

ph2 <- with(imp.tgce, coxph(Surv(t, d) ~ age + figo + grade + histol + ascites + ps + resdis +
log_ca125 + log_alp))
  # 補完後のデータセット（M組）の解析；Cox回帰

pool2 <- pool(ph2)
round(summary(pool2),3)
  # M組の解析結果の統合と結果の表示
```

75

Imputed Datasetの内訳

```
Multiply imputed data set
Call:
mice(data = tgce, m = 200, predictorMatrix = predmt1, seed = 34871)
Number of multiple imputations:  200
Missing cells per column:
    patno       age      figo     grade    histol   ascites        ps    resdis log_ca125   log_alp         d         t
        0         0        21       139         0        65       511        81       440       396         0         0
Imputation methods:
    patno       age      figo     grade    histol   ascites        ps    resdis log_ca125   log_alp         d         t
       ""        "" "polyreg" "polyreg"        ""  "logreg" "polyreg" "polyreg"     "pmm"     "pmm"        ""        ""
VisitSequence:
     figo     grade   ascites        ps    resdis log_ca125   log_alp
        3         4         6         7         8         9        10
PredictorMatrix:
          patno age figo grade histol ascites ps resdis log_ca125 log_alp d t
patno         0   0    0     0      0       0  0      0         0       0 0 0
age           0   0    0     0      0       0  0      0         0       0 0 0
figo          0   1    0     1      1       1  1      1         1       1 0 0
grade         0   1    1     0      1       1  1      1         1       1 0 0
histol        0   0    0     0      0       0  0      0         0       0 0 0
ascites       0   1    1     1      1       0  1      1         1       1 0 0
ps            0   1    1     1      1       1  0      1         1       1 0 0
resdis        0   1    1     1      1       1  1      0         1       1 0 0
log_ca125     0   1    1     1      1       1  1      1         0       1 0 0
log_alp       0   1    1     1      1       1  1      1         1       0 0 0
d             0   0    0     0      0       0  0      0         0       0 0 0
t             0   0    0     0      0       0  0      0         0       0 0 0
Random generator seed value:  34871
```

76

代入値の生成方法のオプション（R: mice）

Method	Description	Scale type	Default
`pmm`	Predictive mean matching	numeric	Y
`norm`	Bayesian linear regression	numeric	
`norm.nob`	Linear regression, non-Bayesian	numeric	
`mean`	Unconditional mean imputation	numeric	
`2L.norm`	Two-level linear model	numeric	
`logreg`	Logistic regression	factor, 2 levels	Y
`polyreg`	Multinomial logit model	factor, >2 levels	Y
`polr`	Ordered logit model	ordered, >2 levels	Y
`lda`	Linear discriminant analysis	factor	
`sample`	Random sample from the observed data	any	

Table 1: Built-in univariate imputation techniques. The techniques are coded as functions named `mice.impute.pmm()`, and so on.

van Buuren and Groothuis-Oudshoorn (2011)

77

Original Dataset

	patno	age	figo	grade	histol	ascites	ps	resdis	log_ca125	log_alp	d	t
1	1	45.67830	I	<NA>	Mucinous	present	0	<2 cm	4.852030	3.737670	0	7.1485284
2	2	70.10815	III	moderately differentiated	Serous papillary	absent	0	>5 cm	NA	4.330733	1	1.1690623
3	3	82.65572	III	poorly differentiated	Mesonephroid (clear cell)	present	<NA>	<2 cm	NA	NA	1	2.2532512
4	4	46.98426	III	poorly differentiated	Adenocarcinoma	absent	1	<2 cm	2.197225	4.234107	1	2.3983573
5	5	65.72485	IV	moderately differentiated	Serous papillary	present	0	>5 cm	NA	4.248495	1	1.4510609
6	6	38.33265	I	well differentiated	Serous papillary	present	<NA>	<2 cm	NA	NA	0	9.4565366
7	7	51.10746	III	poorly differentiated	Serous papillary	present	1	2-5 cm	5.347107	4.663439	1	1.0239562
8	8	63.72074	III	poorly differentiated	Serous papillary	absent	1	>5 cm	7.432484	5.187386	1	0.5585216
9	9	44.06845	I	poorly differentiated	Mucinous	absent	0	<2 cm	3.367296	4.317488	0	11.0171116
10	10	42.57358	III	poorly differentiated	Serous papillary	present	<NA>	>5 cm	NA	NA	1	2.3791923
11	11	60.49007	I	well differentiated	Endometrioid	present	0	<2 cm	NA	5.411646	1	13.2621492
12	12	56.50103	I	well differentiated	Endometrioid	present	1	<2 cm	2.944439	4.418840	1	6.9103354
13	13	67.51814	II	poorly differentiated	Mucinous	absent	0	<2 cm	3.637586	4.682131	1	4.5229295
14	14	57.42368	II	poorly differentiated	Serous papillary	absent	1	2-5 cm	7.989561	4.174388	1	1.1033539
15	15	45.28405	II	well differentiated	Serous papillary	present	0	<2 cm	4.697749	4.430817	1	1.5633128
16	16	71.59753	<NA>	<NA>	Adenocarcinoma	present	3to4	>5 cm	NA	4.574711	1	0.4982888
17	17	64.79398	I	moderately differentiated	Mucinous	present	<NA>	<2 cm	5.209486	NA	1	9.3032170
18	18	67.26900	III	moderately differentiated	Serous papillary	present	1	>5 cm	3.970292	4.189655	1	2.6967830
19	19	48.37235	III	poorly differentiated	Mucinous	absent	1	2-5 cm	5.945421	4.094345	1	1.7138946
20	20	57.04038	II	poorly differentiated	Serous papillary	present	0	<2 cm	NA	4.605170	1	5.3661875
21	21	59.67693	I	well differentiated	Endometrioid	absent	<NA>	<2 cm	NA	NA	0	10.4613279
22	22	43.38398	II	poorly differentiated	Serous papillary	absent	0	<2 cm	3.931826	4.488636	0	11.9041752
23	23	55.67967	I	moderately differentiated	Endometrioid	absent	0	<2 cm	2.484907	4.488636	1	0.8788501
24	24	76.51746	I	poorly differentiated	Serous papillary	absent	0	<2 cm	6.836259	NA	1	2.1574264
25	25	60.61328	I	well differentiated	Serous papillary	absent	<NA>	<2 cm	4.787492	NA	1	2.8966461

78

The 5th Imputed Dataset

	patno	age	figo	grade	histol	ascites	ps	resdis	log_ca125	log_alp	d	t
1	1	45.67830	I	moderately differentiated	Mucinous	present	0	<2 cm	4.852030	3.737670	0	7.14852841
2	2	70.10815	III	moderately differentiated	Serous papillary	absent	0	>5 cm	5.529429	4.330733	1	1.16906229
3	3	82.65572	III	poorly differentiated	Mesonephroid (clear cell)	present	1	<2 cm	4.158883	4.595120	1	2.25325120
4	4	46.98426	III	poorly differentiated	Adenocarcinoma	absent	1	<2 cm	2.197225	4.234107	1	2.39835729
5	5	65.72485	IV	moderately differentiated	Serous papillary	present	0	>5 cm	3.258096	4.248495	1	1.45106092
6	6	38.33265	I	well differentiated	Serous papillary	present	0	<2 cm	4.204693	4.189655	0	9.45653662
7	7	51.10746	III	poorly differentiated	Serous papillary	present	1	2-5 cm	5.347107	4.663439	1	1.02395619
8	8	63.72074	III	poorly differentiated	Serous papillary	absent	1	>5 cm	7.432484	5.187386	1	0.55852156
9	9	44.06845	I	poorly differentiated	Mucinous	absent	0	<2 cm	3.367296	4.317488	0	11.01711157
10	10	42.57358	III	poorly differentiated	Serous papillary	present	0	>5 cm	5.455321	4.488636	1	2.37919233
11	11	60.49007	I	well differentiated	Endometrioid	present	0	<2 cm	2.701361	5.411646	1	13.26214921
12	12	56.50103	I	well differentiated	Endometrioid	present	1	<2 cm	2.944439	4.418840	1	6.91033539
13	13	67.51814	II	poorly differentiated	Mucinous	absent	0	<2 cm	3.637586	4.682131	1	4.52292950
14	14	57.42368	II	poorly differentiated	Serous papillary	absent	1	2-5 cm	7.989561	4.174388	1	1.10335387
15	15	45.28405	II	well differentiated	Serous papillary	present	0	<2 cm	4.697749	4.430817	1	1.56331280
16	16	71.59753	IV	poorly differentiated	Adenocarcinoma	present	3to4	>5 cm	8.187577	4.574711	1	0.49828884
17	17	64.79398	I	moderately differentiated	Mucinous	present	0	<2 cm	5.209486	4.927254	1	9.30321697
18	18	67.26900	III	moderately differentiated	Serous papillary	present	1	>5 cm	3.970292	4.189655	1	2.69678302
19	19	48.37235	III	poorly differentiated	Mucinous	absent	1	2-5 cm	5.945421	4.094345	1	1.71389459
20	20	57.04038	II	poorly differentiated	Serous papillary	present	0	<2 cm	4.605170	4.605170	1	5.36618754
21	21	59.67693	I	well differentiated	Endometrioid	absent	0	<2 cm	4.127134	5.323010	0	10.46132786
22	22	43.38398	II	poorly differentiated	Serous papillary	absent	0	<2 cm	3.931826	4.488636	0	11.90417522
23	23	55.67967	I	moderately differentiated	Endometrioid	absent	0	<2 cm	2.484907	4.488636	1	0.87885010
24	24	76.51746	I	poorly differentiated	Serous papillary	absent	0	<2 cm	6.836259	5.780744	1	2.15742642
25	25	60.61328	I	well differentiated	Serous papillary	absent	0	<2 cm	4.787492	5.068904	1	2.89664613

79

Cox Regression for the 5th Imputed Dataset

```
Call:
coxph(formula = Surv(t, d) ~ age + figo + grade + histol + ascites + ps + resdis + log_ca125 + log_alp)

               coef exp(coef) se(coef)      z        p
age         0.02315   1.02342  0.00355  6.51 7.3e-11
figo2       0.68788   1.98948  0.16169  4.25 2.1e-05
figo3       1.19456   3.30211  0.13852  8.62 < 2e-16
figo4       1.23582   3.44119  0.16746  7.38 1.6e-13
grade2      0.24470   1.27724  0.15931  1.54  0.1245
grade3      0.27934   1.32226  0.15447  1.81  0.0705
histol2     0.23577   1.26588  0.19404  1.22  0.2244
histol3    -0.06239   0.93952  0.15968 -0.39  0.6960
histol4     0.35199   1.42190  0.13625  2.58  0.0098
histol5     0.82749   2.28757  0.18033  4.59 4.5e-06
histol6    -0.21044   0.81023  0.10158 -2.07  0.0383
histol7     0.22521   1.25259  0.24151  0.93  0.3511
ascites2    0.37842   1.45997  0.08382  4.51 6.3e-06
ps2         0.09476   1.09939  0.08883  1.07  0.2861
ps3         0.22873   1.25701  0.11694  1.96  0.0505
ps4         0.67439   1.96284  0.15084  4.47 7.8e-06
resdis2    -0.14265   0.86706  0.10427 -1.37  0.1713
resdis3    -0.65520   0.51934  0.10645 -6.15 7.5e-10
log_ca125   0.03206   1.03258  0.02736  1.17  0.2413
log_alp     0.37301   1.45210  0.07997  4.66 3.1e-06

Likelihood ratio test=744  on 20 df, p=0
n= 1189, number of events= 842
```

欠測値をImputeしたことによって、
すべてのデータを用いた解析を行うことができる

80

Cox Regression for the 25th Imputed Dataset

```
Call:
coxph(formula = Surv(t, d) ~ age + figo + grade + histol + ascites + ps + resdis + log_ca125 + log_alp)

               coef exp(coef) se(coef)     z       p
age        0.02261   1.02287  0.00354  6.38 1.7e-10
figo2      0.56968   1.76770  0.16207  3.51 0.00044
figo3      1.23025   3.42208  0.13846  8.89 < 2e-16
figo4      1.31465   3.72345  0.16522  7.96 1.8e-15
grade2     0.51168   1.66809  0.16442  3.11 0.00186
grade3     0.54567   1.72576  0.15930  3.43 0.00061
histol2    0.20187   1.22369  0.19102  1.06 0.29060
histol3   -0.00814   0.99189  0.15886 -0.05 0.95914
histol4    0.34285   1.40895  0.13779  2.49 0.01284
histol5    0.82787   2.28843  0.18256  4.53 5.8e-06
histol6   -0.16791   0.84543  0.10251 -1.64 0.10143
histol7    0.48894   1.63059  0.23860  2.05 0.04044
ascites2   0.30533   1.35708  0.08322  3.67 0.00024
ps2        0.04053   1.04136  0.08671  0.47 0.64021
ps3        0.14208   1.15267  0.12043  1.18 0.23808
ps4        0.68841   1.99054  0.14888  4.62 3.8e-06
resdis2   -0.16324   0.84939  0.10225 -1.60 0.11039
resdis3   -0.64448   0.52494  0.10337 -6.23 4.5e-10
log_ca125  0.01870   1.01888  0.02752  0.68 0.49687
log_alp    0.38169   1.46476  0.07808  4.89 1.0e-06

Likelihood ratio test=738  on 20 df, p=0
n= 1189, number of events= 842
```

81

Cox Regression for the 50th Imputed Dataset

```
Call:
coxph(formula = Surv(t, d) ~ age + figo + grade + histol + ascites + ps + resdis + log_ca125 + log_alp)

               coef exp(coef) se(coef)     z       p
age        0.02311   1.02338  0.00354  6.54 6.4e-11
figo2      0.64665   1.90913  0.16135  4.01 6.1e-05
figo3      1.20136   3.32463  0.13920  8.63 < 2e-16
figo4      1.23283   3.43093  0.16710  7.38 1.6e-13
grade2     0.32350   1.38195  0.16105  2.01 0.04457
grade3     0.45357   1.57392  0.15582  2.91 0.00361
histol2    0.11690   1.12401  0.19141  0.61 0.54137
histol3   -0.03584   0.96480  0.16119 -0.22 0.82405
histol4    0.28343   1.32768  0.13830  2.05 0.04042
histol5    0.76061   2.13958  0.18172  4.19 2.8e-05
histol6   -0.23698   0.78901  0.10187 -2.33 0.02000
histol7    0.35975   1.43297  0.24124  1.49 0.13589
ascites2   0.33703   1.40078  0.08308  4.06 5.0e-05
ps2        0.10642   1.11229  0.08629  1.23 0.21748
ps3        0.23393   1.26355  0.11923  1.96 0.04976
ps4        0.89309   2.44266  0.15248  5.86 4.7e-09
resdis2   -0.10031   0.90456  0.10355 -0.97 0.33268
resdis3   -0.67208   0.51064  0.10298 -6.53 6.7e-11
log_ca125  0.01057   1.01063  0.02702  0.39 0.69551
log_alp    0.26331   1.30123  0.07178  3.67 0.00024

Likelihood ratio test=739  on 20 df, p=0
n= 1189, number of events= 842
```

82

補完回数 M の設定

- 初期の教科書では、補完回数 M は3~5回で十分であるとされてきた（例えば、Rubin (1987)）
- もともと多重補完法が提案された1970~80年代では、十分な性能を持つ計算機がなく、Mを大きくしたもとでの計算は、現実問題として困難だという背景もあった
- Mを大きくとるほど、信頼区間・P値の近似精度は高くなる
  - 小さすぎると、近似精度が不十分である可能性も
- 最近の文献では、正確な推定・検定を行うために、100～1000回のオーダーにとることを勧めているものも多い

Carpenter and Kenward (2013), Royston and White (2011)

83

R example: MICE Output

```
> round(summary(pool2),3)
             est    se      t         df Pr(>|t|)  lo 95  hi 95 nmis   fmi lambda
age        0.024 0.004  6.436  77589.144    0.000  0.016  0.031    0 0.049  0.049
figo2      0.633 0.165  3.830  79236.415    0.000  0.309  0.956   NA 0.048  0.048
figo3      1.184 0.143  8.278  48394.035    0.000  0.904  1.465   NA 0.062  0.062
figo4      1.246 0.175  7.129  23140.643    0.000  0.904  1.589   NA 0.092  0.092
grade2     0.384 0.173  2.218  14749.187    0.027  0.045  0.723   NA 0.115  0.115
grade3     0.437 0.167  2.608  16488.140    0.009  0.109  0.765   NA 0.109  0.109
histol2    0.181 0.204  0.888  16983.350    0.375 -0.219  0.581   NA 0.107  0.107
histol3   -0.031 0.165 -0.188  59966.434    0.851 -0.354  0.292   NA 0.056  0.056
histol4    0.325 0.142  2.284  50458.739    0.022  0.046  0.603   NA 0.061  0.061
histol5    0.802 0.190  4.214  29427.975    0.000  0.429  1.174   NA 0.081  0.081
histol6   -0.206 0.104 -1.975 104717.958    0.048 -0.410 -0.002   NA 0.041  0.041
histol7    0.333 0.267  1.248   5984.292    0.212 -0.190  0.856   NA 0.182  0.182
ascites2   0.334 0.089  3.750  16858.463    0.000  0.159  0.508   NA 0.108  0.108
ps2        0.106 0.101  1.048   3096.892    0.295 -0.092  0.304   NA 0.253  0.253
ps3        0.207 0.144  1.443   2083.436    0.149 -0.074  0.489   NA 0.309  0.309
ps4        0.628 0.209  3.001    949.924    0.003  0.217  1.039   NA 0.458  0.457
resdis2   -0.116 0.112 -1.035  11215.134    0.301 -0.336  0.104   NA 0.132  0.132
resdis3   -0.650 0.112 -5.801   9575.774    0.000 -0.870 -0.431   NA 0.144  0.143
log_ca125  0.031 0.031  1.008   2997.766    0.314 -0.030  0.092  440 0.258  0.257
log_alp    0.331 0.096  3.428   1675.250    0.001  0.142  0.520  396 0.345  0.344
```

84

解析結果の比較①

	CCA (n=362, # deaths=248)				MICE (n=1189, # deaths=842)			
	HR	95%CI		P-value	HR	95%CI		P-value
Age (years)	1.02	1.00	1.03	0.011	1.02	1.02	1.03	< 0.001
FIGO stage								
I	1.00				1.00			
II	2.39	1.14	5.00	0.021	1.88	1.36	2.60	< 0.001
III	4.01	2.12	7.59	< 0.001	3.27	2.47	4.33	< 0.001
IV	5.83	2.81	12.09	< 0.001	3.48	2.47	4.90	< 0.001
Grade								
I	1.00				1.00			
II	1.28	0.65	2.52	0.471	1.47	1.05	2.06	0.027
III	1.13	0.59	2.13	0.717	1.55	1.11	2.15	0.009

†CCA: Complete-Case Analysis, MICE: Multiple Imputation by Chained Equation

85

解析結果の比較②

	CCA (n=362, # deaths=248)				MICE (n=1189, # deaths=842)			
	HR	95%CI		P-value	HR	95%CI		P-value
Histology								
Serous papillary	1.00				1.00			
Adenocarcinoma	1.54	0.37	6.36	0.554	1.20	0.80	1.79	0.375
Endometrioid	0.96	0.52	1.78	0.901	0.97	0.70	1.34	0.851
Clear cell	1.71	1.00	2.93	0.051	1.38	1.05	1.83	0.022
Mixed mesodermal	3.99	1.78	8.91	0.001	2.23	1.54	3.24	< 0.001
Mucinous	0.77	0.55	1.08	0.126	0.81	0.66	1.00	0.048
Undifferentiated	1.43	0.60	3.41	0.421	1.40	0.83	2.35	0.212
Ascites								
Absence	1.00				1.00			
Presence	1.35	0.98	1.86	0.069	1.40	1.17	1.66	< 0.001

†CCA: Complete-Case Analysis, MICE: Multiple Imputation by Chained Equation

86

解析結果の比較③

	CCA (n=362, # deaths=248)				MICE (n=1189, # deaths=842)			
	HR	95%CI		P-value	HR	95%CI		P-value
Performance status								
0	1.00				1.00			
1	1.16	0.86	1.57	0.325	1.11	0.91	1.35	0.295
2	0.97	0.62	1.50	0.885	1.23	0.93	1.63	0.149
3+4	2.30	0.96	5.48	0.061	1.87	1.24	2.83	0.003
Residual disease								
> 5 cm	1.00				1.00			
2-5 cm	0.97	0.67	1.41	0.880	0.89	0.71	1.11	0.301
< 2 cm	0.57	0.39	0.82	0.003	0.52	0.42	0.65	< 0.001
Log CA125	1.06	0.97	1.17	0.205	1.03	0.97	1.10	0.314
Log alkaline phos.	1.61	1.14	2.27	0.006	1.39	1.15	1.68	0.001

†CCA: Complete-Case Analysis, MICE: Multiple Imputation by Chained Equation

87

Some Remarks

- 近年の研究により、理論的妥当性が厳密に保証される条件はやや限定的であることが示されている
- ただ、多くの性能評価研究で、かなり広範な条件下で良好な成績が示されている（少なくとも、実践的には Favorable or Acceptableであることが多い）
- また、複数の変数にまたがって欠測が起こるような一般的な条件下で、汎用的なソフトウェアが充実している数少ない方法であり、また今後、実践でもますます普及していくことが予想される

White et al. (2011), Royston and White (2011)

88

Contexts

- 欠測データの統計解析：医学研究における応用から
  - National Research Councilによる調査報告書の推奨事項などを中心に、欠測データの統計解析の実践的な方法論の基礎について、午前中の復習も兼ねて解説を行います
- Case Study 1: Multiple Imputation by Chained Equation (MICE)
  - 複数の変数に非単調に起こった欠測を扱うには？？
- Case Study 2: 脱落を伴う経時測定データの統計解析
  - 非ランダムなメカニズムで起こる脱落によるデータの欠測をどのように扱えばよいのか？？

89

ACTG193A試験

- AIDS Clinical Trials Group (ACTG) によって、1990年代に米国で行われた臨床試験（Henry et al., 1993)
- 1313人の参加者に対して、4種類の逆転写酵素阻害剤の併用療法・交替療法をランダムに割り付け
- ここでは、Zidovudine, Didanozine, Nevirapineの3剤併用群（n=300）とそれ以外の2剤併用・交替療法の3群（n=979) を比較することとする

90

経時測定データにおける脱落（欠測）

- 治療の有効性を評価するアウトカムを「32週時点のCD4細胞数」とする
- 長期間の追跡を行い、治療効果の有効性を評価するような試験では、当然ながら、追跡途中での脱落や、間歇的な欠測が起こる
- ここでは、単調な欠測（脱落のみ）を考えることとする

† 実際には、主要評価項目は、”Time to death” であったが、ここでは、副次的評価項目の1つを解析対象として考えている。
* CD4細胞数は、免疫系の情報伝達に関与する白血球の一種であり、HIVの感染によってその数が減少していく。疾患の進行度を測る指標として用いられる。

91

log(CD4+1) のデータの時点ごとの要約

	追跡開始時点	8週目	16週目	24週目	32週目
3剤併用群（n=330）					
欠測	–	22	50	84	122
（%）	–	(6.7%)	(15.2%)	(25.5%)	(37.0%)
平均	2.83	3.27	3.25	3.06	2.99
標準偏差	0.96	1.13	1.17	1.14	1.23
Otherwise（n=979）					
欠測	–	100	164	263	413
（%）	–	(10.2%)	(16.8%)	(26.9%)	(42.2%)
平均	2.94	2.93	2.86	2.73	2.67
標準偏差	0.90	1.13	1.10	1.07	1.06
P値（t検定）	–	< 0.001	< 0.001	< 0.001	< 0.001

92

経時測定データにおける脱落（欠測）

- 32週時点での欠測は、3剤併用群で122人 (37.0%)，それ以外の3群で413人 (42.2%) も起こっている
- 32週時点での Available Data によって、群間差の検定をt検定で行うと、P < 0.001 となるが…
- 仮に、それぞれの群が異なるメカニズムで欠測を起こしているのだとすると、治療のランダム割り付けによる恩恵は失われ、群間の治療効果の評価にはバイアスが生じることとなる

93

欠測メカニズムは？？

- 欠測が「何の理由もなく起こる」ということはほとんどあり得ず、「欠測を起こした対象者には、欠測を起こすだけのなんらかの理由がある」ことが一般的である
- NRC Reportで議論されていた通り、症状の進行によって、試験治療を継続できなくなった患者や、副作用によって試験治療が中止される患者、あるいは、追跡終了時点までになくなった患者など、さまざまな原因によって欠測が起こったことが考えられる

94

経時測定データの解析

- 入門レベルの統計学の教科書で解説される手法のほとんどは、測定されたデータがすべて「独立」であることを前提としている
- ACTG193A試験のように、同一個人に対して、繰り返し測定されるデータは、一般的に独立と見なすことはできない
- これらの経時測定データの解析では、妥当な結論を得るために、繰り返し測定されるデータの時点間の相関を適切に考慮した解析が必要となる

95

Notation

- 結果変数 $\boldsymbol{y}_i = (y_{i1}, \dots, y_{iT})^t \quad (i = 1,2, \dots, n)$
  - 指数型分布（正規分布, 2項分布, ポアソン分布など）に従うものとする
- 共変量行列 $\boldsymbol{X}_i = \left(\boldsymbol{x}_{i1}^t, \dots, \boldsymbol{x}_{in}^t\right)^t, \boldsymbol{W}_i = \left(\boldsymbol{w}_{i1}^t, \dots, \boldsymbol{w}_{in}^t\right)^t$
- 回帰関数 $\boldsymbol{\mu}_i = E(\boldsymbol{y}_i | \boldsymbol{X}_i)$
  - $\boldsymbol{X}_i, \boldsymbol{W}_i$ は、回帰関数に含まれる共変量かどうかによって分類される

96

周辺モデル（Marginal Model）

- 一般化線形モデルの多変量モデルへの拡張
  - $g(\boldsymbol{\mu}_i) = \left(g(\boldsymbol{\mu}_{i1}), \ldots, g(\boldsymbol{\mu}_{iT})\right)^t = \boldsymbol{X}_i\boldsymbol{\beta}$
  - $g(.)$ はリンク関数
- 結果変数の時点間の分散共分散行列
  - $\boldsymbol{\Sigma}_i = V(\boldsymbol{y}_i|\boldsymbol{X}_i) = \phi\boldsymbol{A}_i(\boldsymbol{\beta})^{1/2}\boldsymbol{C}_i(\boldsymbol{\alpha})\boldsymbol{A}_i(\boldsymbol{\beta})^{1/2}$
    - $\boldsymbol{A}_i(\boldsymbol{\beta})$: 分散関数 $a(\mu_{i1}), \ldots, a(\mu_{iT})$ を対角成分に持つ対角行列
    - $\boldsymbol{C}_i(\boldsymbol{\alpha})$: 結果変数の時点間の相関係数行列

Liang and Zeger (1986), Diggle et al. (2002)

97

一般化推定方程式（GEE）

$$U_{GEE}(\beta) = \sum_{i=1}^{n} \boldsymbol{D}_i^t \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_i) = 0$$

Liang and Zeger (1986)

- $\boldsymbol{D}_i = \boldsymbol{D}_i(\boldsymbol{\beta}) = (\partial\boldsymbol{\mu}_i/\partial\boldsymbol{\beta})^t$
- 仮に、分散関数 $\boldsymbol{A}_i(\boldsymbol{\beta})$ や時点間の相関係数行列 $\boldsymbol{C}_i(\boldsymbol{\alpha})$ が誤特定されていたとしても、推定関数の不偏性は成り立つ
  - 推定量の一致性は成り立つ
  - $\boldsymbol{C}_i(\alpha)$ は、その意味で作業相関行列と呼ばれる
- Wedderburn (1974) の擬似尤度法の一般化と考えることができ、1次と2次のモーメントについてのみパラメトリックな仮定を置いたセミパラメトリック推測の方法となる
- Generalized Estimating Equation (GEE) と呼ばれる

98

多変量正規線形回帰モデル

- 多変量正規回帰モデル
  - $\boldsymbol{\mu}_i = \boldsymbol{X}_i\boldsymbol{\beta}$
  - $\boldsymbol{y}_i$ の分布型が正規分布、リンク関数が恒等リンク（identity link）であるモデル
- 結果変数間の作業相関構造については、独立 (IND)、交換可能 (EXC)、1次の自己回帰 (AR(1))、無構造 (UN) などのモデルがある
  - 誤特定されていても一致性は成り立つが、真の相関構造に近いモデルほど、$\boldsymbol{\beta}$ の推定精度は向上する

99

補足：作業相関行列のモデル（3時点の場合）

$$\boldsymbol{C}_{IND} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{C}_{EXC} = \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$

$$\boldsymbol{C}_{AR(1)} = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}, \boldsymbol{C}_{UN} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix},$$

† ρは、隣り合った時点間の結果変数の相関係数を表す。
* 交換可能 (exchangeable) は、複合対称 (compound symmetry) ともいう。

Diggle et al. (2002) 100

推定関数への寄与

$$U_{GEE}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \underbrace{\boldsymbol{D}_i^t \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_i)}_{= U_i(\boldsymbol{\beta});\ i\text{番目の対象者の寄与}} = 0$$

- 左辺の基準関数を推定関数（Estimating Function）という
- 完全データのもとでは、n人の対象者が、全員すべて「1」の重みで等しく推定関数に寄与する

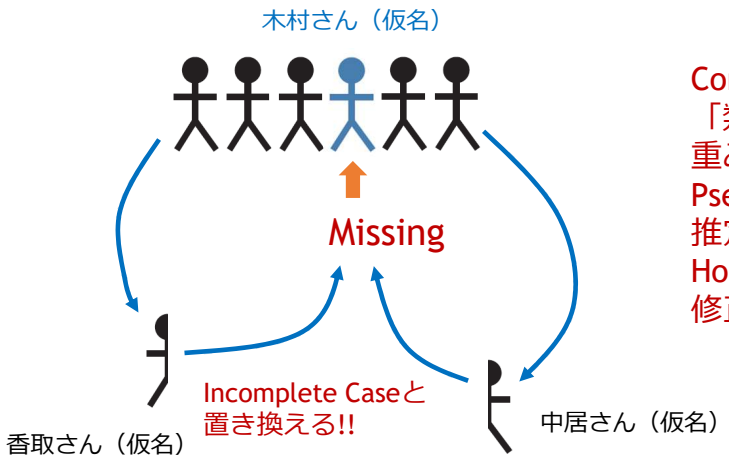
101

Complete-Case Analysis

- ACTG193A試験のように、脱落を伴うデータの場合には？？
- 結果変数・説明変数のいずれか1つでも欠測がある対象者を除外
- 推定関数内での重みは、
  - 完全データが観測されている対象者→「1」
  - 欠測が1つでもある対象者→「0」
- 個人ごとの重みを調整し、「適当な重み」を付与し直すことによって、バイアスをゼロにすることはできるのではないか?

102

IPW法の考え方



Complete Dataが観測された
「類似した情報を持つ対象者」を
重みつき平均してできる、
Pseudo-Obserbationを
推定関数の中に
Hot-dec Imputationして
修正推定関数を構成する!!

103

推定関数の構成方法

- 実際には、欠測を起こした対象者はたくさんおり、そのすべてをHot-dec補完しなくてはいけない
- どの対象者のデータをどの程度の割合で置き換えればよいのか？
- 理論的結果：i番目の対象者の推定関数の寄与 $U_i(\boldsymbol{\beta})$ を「i番目の対象者が完全データが観測される確率」の逆数分だけ複製して、欠測を起こした対象者と置き換えた、修正推定関数によって、$\boldsymbol{\beta}$の妥当な（漸近的にバイアスのない）推定量が得られる！！

Robins et al. (1994)　104

IPW-GEE (subject level)

$$U_{IPWCC}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \frac{r_{iT}}{\pi_{iT}} \{\boldsymbol{D}_i^t \boldsymbol{\Sigma}_i^{-1} (\boldsymbol{y}_i - \boldsymbol{\mu}_i)\} = 0$$

- 最終時点までの完全データが得られた対象者を単位としたIPW推定方程式
  - r_{iT}: 対象者 i が最終時点 t においてデータが観測されるかどうかを表す指示変数
  - π_{iT}: 対象者 i が最終時点 t においてデータが観測される確率

Robins et al. (1994), 高井ら (2016)

Available Data Analysis

- 経時測定データは、個々人ごとに複数の時点のデータが観測されている多変量データとなるため、最終時点までのデータ（完全データ）が得られている対象者だけではなく、途中で脱落を起こした対象者にも、「途中までの利用可能なデータ（Available Data）」が存在する
- 途中までの利用可能なデータを組み込んだ解析のことを、Available Data Analysisという
- MARの仮定のもとでは、尤度に基づく方法では一致推定量が得られる
- GEEは、最尤法とは異なるセミパラメトリック推測の方法なので、MARでも一致性が保証されない！！

高井ら (2016)

IPW-GEE (observation level)

- IPW法は、「完全データが得られた個人」を単位にするのではなく、「時点ごとの観測値」を単位にして重みを定義することもできる
- 対象者 i が時点 t までに脱落を起こさない確率

$$\pi_{it}(\boldsymbol{\alpha}) = \Pr(r_{it} = 1 | \mathcal{F}_{it}) = \prod_{j=1}^{t} \{1 - \lambda_{ij}(\boldsymbol{\alpha})\}$$

- $\lambda_{it}(\boldsymbol{\alpha}) = \Pr\left(r_{it} = 0 \middle| r_{i,t-1} = 1, \mathcal{F}_{it}\right)$
  - r_{it}: 対象者 i が時点 t においてデータが観測されるかどうかを表す指示変数

Robins et al. (1994), 高井ら (2016) 107

IPW-GEE (observation level)

$$U_{IPW}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \boldsymbol{D}_i^t \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Delta}_i(\hat{\boldsymbol{\alpha}})(\boldsymbol{y}_i - g^{-1}(\boldsymbol{X}_i, \boldsymbol{\beta})) = 0$$

- $\boldsymbol{\Delta}_i(\boldsymbol{\alpha})$: $\{r_{i1}/\pi_{i1}(\boldsymbol{\alpha}), \ldots, r_{iT}/\pi_{iT}(\boldsymbol{\alpha})\}$ を対角成分に持つ対角行列
- Available Dataを単位にしたIPW推定方程式
  - ↔ Subject levelのIPW法は、Complete-Caseを単位にしたIPW推定方程式

Robins et al. (1994), 高井ら (2016) 108

欠測確率の推定と正則条件

- 欠測確率 $\lambda_{it}(\boldsymbol{\alpha})$ は、実際には未知なので、観測された共変量のデータ $(\boldsymbol{X}, \boldsymbol{W})$ をもとに、推定する必要がある
  - 観測・欠測の指示変数 r_{it} を結果変数としたロジスティック回帰モデルなどで推定を行う
- $\boldsymbol{\beta}$ の推定の一致性が成り立つための条件は…
  - (i) 欠測確率 $\lambda_{it}(\boldsymbol{\alpha})$ の回帰モデルが正しく特定されていること、(ii) すべての時点の脱落確率が 1 にはならないこと（観測確率が 0 となり、重みが不定となる）、(iii) Sequential-MARが成り立つ

†Sequential-MARは、MARより少し強い条件。詳細は、Robins et al. (1994) をご参照ください。

109

漸近分散の推定量

- 一般的な推定方程式の理論にしたがって、サンドイッチ分散（ロバスト分散）の推定量によって、推定することができる
- ただし、結果変数の回帰モデルと、欠測確率の推定のためのロジスティック回帰モデルの同時方程式から構成されるロバスト分散となるため、式の型はかなり複雑になる
  - Robins et al. (1994) を参照
- SASのPROC GEEなどでは、この同時方程式から得られるロバスト分散の推定値が計算される

110

SAS Example: PROC GEE

```
/* IPW-GEE, subject level */
proc gee data=datalib.&DataName desc plots=histogram;
  class PATIENT TRT SEQ GENDER;
  missmodel SEQ TRT PREV AGE GENDER/ type=sublevel;   /* missingness model */
  model Y = TRT BASVAL SEQ TRT*SEQ / dist=normal;
  repeated subject=PATIENT / within=SEQ corr=un CORRW;
run;

/* IPW-GEE, observation level */
proc gee data=datalib.&DataName desc plots=histogram;
  class PATIENT TRT SEQ GENDER;
  missmodel SEQ TRT PREV AGE GENDER/ type=obslevel;   /* missingness model */
  model Y = TRT BASVAL SEQ TRT*SEQ / dist=normal;
  repeated subject=PATIENT / within=SEQ corr=un CORRW;
run;
```

111

IPW-GEE (subject-level)

GEE Model Information	
Correlation Structure	Unstructured
Within-Subject Effect	SEQ (4 levels)
Subject Effect	PATIENT (1187 levels)
Number of Clusters	1187
Clusters With Missing Values	413
Correlation Matrix Dimension	4
Maximum Cluster Size	4
Minimum Cluster Size	1

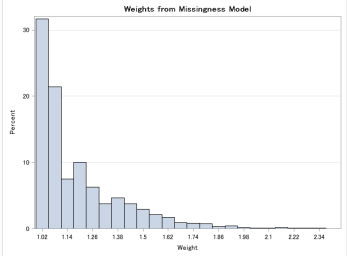
Working Correlation Matrix				
	Obs 1	Obs 2	Obs 3	Obs 4
Obs 1	1.0000	0.3676	0.3378	0.3470
Obs 2	0.3676	1.0000	0.5914	0.5524
Obs 3	0.3378	0.5914	1.0000	0.5875
Obs 4	0.3470	0.5524	0.5875	1.0000

Parameter Estimates for Response Model
with Empirical Standard Error

Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			0.6925	0.0968	0.5028	0.8822	7.15	<.0001
TRT	0		−0.3655	0.0775	−0.5174	−0.2136	−4.72	<.0001
TRT	1		0.0000	0.0000	0.0000	0.0000	.	.
BASVAL			1.7495	0.0574	1.6369	1.8620	30.47	<.0001
SEQ	8		0.3839	0.0621	0.2621	0.5056	6.18	<.0001
SEQ	16		0.3325	0.0629	0.2101	0.4568	5.30	<.0001
SEQ	24		0.1194	0.0610	−0.0000	0.2389	1.96	0.0500
SEQ	32		0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	0	8	−0.0341	0.0751	−0.1813	0.1131	−0.45	0.6498
TRT*SEQ	0	16	−0.0891	0.0734	−0.2329	0.0547	−1.21	0.2246
TRT*SEQ	0	24	−0.0243	0.0714	−0.1643	0.1156	−0.34	0.7333
TRT*SEQ	0	32	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	8	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	16	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	24	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	32	0.0000	0.0000	0.0000	0.0000	.	.

Parameter Estimates for Missingness Model

Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept		0.8266	0.3013	0.2362	1.4170	2.74	0.0061
SEQ	8	0.0000	0.0000	0.0000	0.0000	.	.
SEQ	16	1.0229	0.1366	0.7553	1.2906	7.49	<.0001
SEQ	24	0.5361	0.1242	0.2927	0.7795	4.32	<.0001
SEQ	32	0.0000	0.0000	0.0000	0.0000	.	.
TRT	0	0.0064	0.1257	−0.2400	0.2527	0.05	0.9594
TRT	1	0.0000	0.0000	0.0000	0.0000	.	.
PREV		0.6668	0.1086	0.4540	0.8796	6.14	<.0001
AGE		−0.0044	0.0066	−0.0173	0.0086	−0.66	0.5062
GENDER	F	−0.3109	0.1586	−0.6219	0.0000	−1.96	0.0500
GENDER	M	0.0000	0.0000	0.0000	0.0000	.	.



112

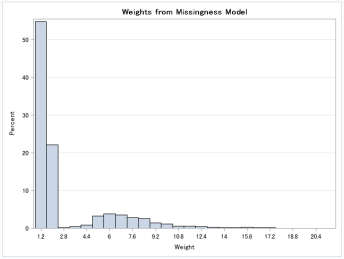
IPW-GEE (observation-level)

GEE Model Information	
Correlation Structure	Unstructured
Within-Subject Effect	SEQ (4 levels)
Subject Effect	PATIENT (1187 levels)
Number of Clusters	1187
Clusters With Missing Values	413
Correlation Matrix Dimension	4
Maximum Cluster Size	4
Minimum Cluster Size	1

Working Correlation Matrix				
	Obs 1	Obs 2	Obs 3	Obs 4
Obs 1	1.0000	0.3464	0.2344	0.2027
Obs 2	0.3464	1.0000	0.5103	0.3246
Obs 3	0.2344	0.5103	1.0000	0.3444
Obs 4	0.2027	0.3246	0.3444	1.0000

Parameter Estimates for Response Model with Empirical Standard Error							
Parameter			Estimate	Standard Error	95% Confidence Limits	Z	Pr > |Z|
Intercept			0.7670	0.1137	0.5441	0.9899	6.75	<.0001
TRT	0		−0.4605	0.0744	−0.6063	−0.3148	−6.19	<.0001
TRT	1		0.0000	0.0000	0.0000	0.0000	.	.
BASVAL			1.7340	0.0674	1.6019	1.8661	25.73	<.0001
SEQ	8		0.2923	0.0719	0.1514	0.4331	4.07	<.0001
SEQ	16		0.3540	0.0828	0.1917	0.5162	4.28	<.0001
SEQ	24		0.1472	0.0707	0.0085	0.2858	2.08	0.0375
SEQ	32		0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	0	8	0.0322	0.0848	−0.1340	0.1983	0.38	0.7042
TRT*SEQ	0	16	−0.1698	0.0932	−0.3520	0.0135	−1.82	0.0694
TRT*SEQ	0	24	−0.0380	0.0833	−0.2012	0.1252	−0.46	0.6478
TRT*SEQ	0	32	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	8	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	16	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	24	0.0000	0.0000	0.0000	0.0000	.	.
TRT*SEQ	1	32	0.0000	0.0000	0.0000	0.0000	.	.

Parameter Estimates for Missingness Model							
Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept		0.8266	0.3013	0.2362	1.4170	2.74	0.0061
SEQ	8	0.0000	0.0000	0.0000	0.0000	.	.
SEQ	16	1.0229	0.1366	0.7553	1.2906	7.49	<.0001
SEQ	24	0.5361	0.1242	0.2927	0.7795	4.32	<.0001
SEQ	32	0.0000	0.0000	0.0000	0.0000	.	.
TRT	0	0.0064	0.1257	−0.2400	0.2527	0.05	0.9594
TRT	1	0.0000	0.0000	0.0000	0.0000	.	.
PREV		0.6668	0.1086	0.4540	0.8796	6.14	<.0001
AGE		−0.0044	0.0066	−0.0173	0.0086	−0.66	0.5082
GENDER	F	−0.3169	0.1586	−0.6219	0.0000	−1.96	0.0500
GENDER	M	0.0000	0.0000	0.0000	0.0000	.	.



113

変量効果モデル

- 結果変数の時点間の相関を、変量効果によってモデル化する
- 個人ごとの変量効果 $\boldsymbol{\eta}_i$ が与えられたもとで、$\{y_{it}\}$ が条件付き独立となるという仮定を置く
  - $g\{E(\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{\eta}_i)\} = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\eta}_i$
  - $\boldsymbol{\eta}_i$はなんらかの確率分布（一般的には、正規分布）に従う変量効果
  - $\boldsymbol{Z}_i$: 個人ごとの変量効果に対応する共変量行列
- 一般化線形混合モデル（generalized linear mixed model; GLMM) と呼ばれる

Diggle et al. (2002), 高井ら (2016)

114

GLMMのパラメータ推定

- 最尤法（maximum likelihood method）
  - 変量効果 $\boldsymbol{\eta}_i$ を積分した周辺尤度を最大化するパラメータを推定値とする
- 制限付き最尤法（restricted maximum likelihood; REML）
  - 最尤推定量の有限標本下でのバイアスを軽減するために、パラメータを探索する空間を制限した最尤法
  - 多くの性能評価研究において、推定量のバイアスや平均二乗誤差などは、REML推定量のほうが優れているという結果が報告されている
- 詳細については、Verbeke and Molenbergh (2000), McCulloch et al. (2008) などを参照ください

115

MARのもとでの妥当性

- MARのもとで起こった欠測に対して、GLMMの最尤法・制限付き最尤法による推測は、尤度に基づく推測の原理によって、Available Data Analysisで妥当な推測を行うことができる
- ただし、尤度に基づく推測なので、GEEのようなセミパラメトリック推測とは異なり、モデルのパラメトリックな仮定が完全に正しくないと一致性・漸近有効性などは成立しない
- 漸近分散の推定や、パラメータの検定・信頼区間の構成などは、通常の尤度に基づく方法のものをそのまま用いることができる

116

線形混合モデル

- $\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$
  - $\boldsymbol{\eta}_i \sim MVN(0, \boldsymbol{S}), \boldsymbol{\varepsilon}_i \sim MVN(0, \boldsymbol{\Sigma})$
- $\boldsymbol{y}_i$ の分布型が正規分布、リンク関数が恒等リンク（identity link）であるモデル
- 医薬品開発の臨床試験では、連続変数をアウトカムした臨床試験では、LOCFが長らく慣用されてきたが、NRCレポートでその科学的妥当性について疑問が呈されたため、線形混合モデルを用いた MMRM (mixed effect models for repeated measurements) というモデルが普及している

Mallinckrodt et al. (2001)

117

経時測定データの解析への応用

- 線形混合モデルの結果変数の周辺分布は、以下の多変量正規分布で表現することができる
  - $\boldsymbol{y}_i \sim MVN(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{V}_i), \boldsymbol{V}_i = \boldsymbol{Z}_i\boldsymbol{S}\boldsymbol{Z}_i^t + \boldsymbol{\Sigma}$
- 一般的な線形混合モデルでは、前頁のパラメトリックモデルのもとで、$\boldsymbol{S}, \boldsymbol{\Sigma}$ を別々に推定する
- 結果変数の時点間の相関構造は $\boldsymbol{V}_i$ で規定されるが、$\boldsymbol{S}, \boldsymbol{\Sigma}, \boldsymbol{Z}_i$を複合した統計量であり、GEEの作業相関行列（複合対称、自己回帰、無構造など）のように直接的な解釈ができない
- 経時測定データの解析において、より柔軟に時点間の相関構造をモデル化できないか？？

118

MMRM

- MMRMとは、2つの分散共分散パラメータを分離せず、1つにまとめて（$\boldsymbol{V}_i$をパラメータとして）、周辺分布 $\boldsymbol{y}_i \sim MVN(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{V}_i)$ そのものをパラメタライズしたモデル
- 通常の線形混合モデルと同様の回帰関数のモデルを仮定しつつ、分散共分散パラメ–タについては、$\boldsymbol{V}_i$を直接パラメトリックにモデル化することができる
  - 時点間の相関構造を、複合対称、自己回帰、無構造などで直接モデル化することができる
- 主たる関心がある固定効果パラメータ $\boldsymbol{\beta}$ については、通常の線形混合モデルと同様の解釈ができる

Mallinckrodt et al. (2001)

MMRMの推測

- 最尤法・制限付き最尤法によるパラメータ推定は、多変量正規分布の尤度関数を用いて、直接尤度法と同じ要領で行うことができる
  - $\boldsymbol{y}_i \sim MVN(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{V}_i)$
- GEEとは異なり、尤度ベースの推測である
  - 時点間の相関係数行列の構造を誤特定すると $\boldsymbol{\beta}$ の一致性も失われるため、注意！！（保守的に無構造が用いられるのが一般的）
- 検定・信頼区間は、通常の線形混合モデルと同様に構成することができる

Mallinckrodt et al. (2001)

SAS Example: PROC MIXED (MMRM)

```
/* MMRM */
proc mixed data=datalib.&DataName;
  class PATIENT TRT SEQ GENDER;
  model Y = TRT BASVAL SEQ TRT*SEQ/s cl ddfm=kr;
  lsmeans TRT*SEQ;
  repeated SEQ/ type=un subject=PATIENT;
run;
```

121

MMRM: SAS-output

オブザベーション数	
読み込んだオブザベーション	4748
使用されたオブザベーション	4018
使用されなかったオブザベーション	730

反復履歴			
反復	評価	-2 残差対数尤度	基準
0	1	10587.56156330	
1	2	9571.23413076	0.00003151
2	1	9571.19910980	0.00000001

収束基準は満たされました。

共分散パラメータ推定値		
共分散パラメータ	サブジェクト	推定値
UN(1,1)	PATIENT	0.7159
UN(2,1)	PATIENT	0.2963
UN(2,2)	PATIENT	0.8338
UN(3,1)	PATIENT	0.2721
UN(3,2)	PATIENT	0.4957
UN(3,3)	PATIENT	0.8702
UN(4,1)	PATIENT	0.2556
UN(4,2)	PATIENT	0.4593
UN(4,3)	PATIENT	0.4816
UN(4,4)	PATIENT	0.8336

適合度統計量	
-2 残差対数尤度	9571.2
AIC (小さいほどよい)	9591.2
AICC (小さいほどよい)	9591.3
BIC (小さいほどよい)	9642.0

帰無モデルの尤度比検定		
自由度	カイ 2 乗値	Pr > ChiSq
9	1016.36	<.0001

固定効果の解										
効果	TRT	SEQ	推定値	標準誤差	自由度	t 値	Pr > |t|	アルファ	下限	上限
Intercept			0.6965	0.09004	1546	7.62	<.0001	0.05	0.5099	0.8831
TRT	0		-0.4537	0.07003	917	-6.48	<.0001	0.05	-0.5912	-0.3163
TRT	1		0	.	.	.	.	.	.	.
BASVAL			1.7947	0.05307	1209	33.82	<.0001	0.05	1.6906	1.8988
SEQ		8	0.3332	0.06533	958	5.10	<.0001	0.05	0.2050	0.4614
SEQ		16	0.2798	0.05770	838	4.85	<.0001	0.05	0.1666	0.3931
SEQ		24	0.08321	0.05819	821	1.43	0.1531	0.05	-0.03100	0.1974
SEQ		32	0	.	.	.	.	.	.	.
TRT*SEQ	0	8	0.05253	0.07615	963	0.69	0.4905	0.05	-0.09691	0.2020
TRT*SEQ	0	16	-0.00059	0.06731	842	-0.01	0.9930	0.05	-0.1327	0.1315
TRT*SEQ	0	24	0.02528	0.06787	826	0.37	0.7097	0.05	-0.1079	0.1585
TRT*SEQ	0	32	0	.	.	.	.	.	.	.
TRT*SEQ	1	8	0	.	.	.	.	.	.	.
TRT*SEQ	1	16	0	.	.	.	.	.	.	.
TRT*SEQ	1	24	0	.	.	.	.	.	.	.
TRT*SEQ	1	32	0	.	.	.	.	.	.	.

固定効果の Type 3 検定				
効果	分子の自由度	分母の自由度	F 値	Pr > F
TRT	1	1150	80.38	<.0001
BASVAL	1	1209	1143.46	<.0001
SEQ	3	966	39.32	<.0001
TRT*SEQ	3	966	0.26	0.8535

最小 2 乗平均							
効果	TRT	SEQ	推定値	標準誤差	自由度	t 値	Pr > |t|
TRT*SEQ	0	8	2.9391	0.02854	1175	102.99	<.0001
TRT*SEQ	0	16	2.8326	0.03182	1118	89.03	<.0001
TRT*SEQ	0	24	2.6619	0.03397	1031	78.35	<.0001
TRT*SEQ	0	32	2.5534	0.03603	928	70.87	<.0001
TRT*SEQ	1	8	3.3403	0.04826	1177	69.22	<.0001
TRT*SEQ	1	16	3.2870	0.05424	1123	60.61	<.0001
TRT*SEQ	1	24	3.0904	0.05792	1033	53.35	<.0001
TRT*SEQ	1	32	3.0072	0.06004	912	50.08	<.0001

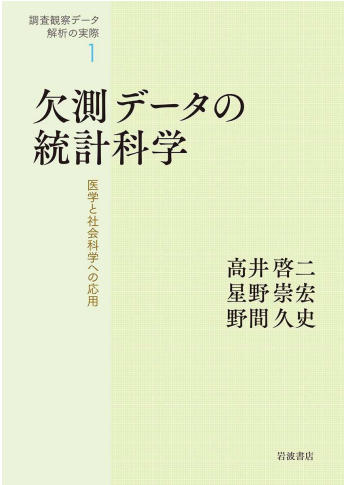
122

まとめ

- 欧米によるガイドライン作成などを受けて、医薬品開発の臨床試験から、アカデミアで行われる臨床研究まで、不完全データの解析に求められる科学的妥当性の要求は高まっている
- 今後、同様の動向は、さまざまな研究領域に拡がっていくことが予想される
- 欠測データの解析方法は、近い将来、調査・実験研究におけるデータ解析のリテラシーのひとつとして求められていくことになると思われる

123

岩波書店：調査観察データ解析の実際



- 欠測データ（不完全データ）の防止や、統計解析上の処理、また、治療-結果の因果推論を行う上でのさまざまな方法論的体系について、最先端の知見をまとめたシリーズです
- 第1巻が、左記の「欠測データの統計科学」となっています
- 本日の講義について、よりAdvancedな内容も含めて、詳細まで解説された教科書となっております

124

Example Dataset & Program

- MNARのもとでの解析プログラムなども含めた、SASプログラム集は、『欠測データの統計科学』（高井啓二，星野崇宏，野間久史著；岩波書店）のホームページからダウンロードできます
  - http://www2.itc.kansai-u.ac.jp/~takai/Book/index.html

125

文献

- Carpenter, J., and Kenward, M. G. (2013). *Multiple Imputation and Its Application*. Chichester: Wiley.
- Clark, T. G., and Altman, D. G. (2003). Developing a prognostic model in the presence of missing data: an ovarian cancer case study. *Journal of Clinical Epidemiology* **56**, 28-37.
- Clark, T. G., Stewart, M. E., Altman, D. G., Gabra, H., and Smyth, J. F. (2001). A prognostic model for ovarian cancer. *British Journal of Cancer* **85**, 944-952.
- Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34**, 187-220.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* **39**, 1-38.
- Diggle, P. J., Heagerty, P. J., Liang, K.-Y., and Zeger, S. L. (2002). *Analysis of Longitudinal Data*, 2nd edition. Oxford: Oxford University Press.
- Henry, K., Erice, A., Tierney, C., *et al.* (1998). A randomized, controlled, double-blind study comparing the survival benefit of four different reverse transcriptase inhibitor therapies (three-drug, two-drug, and alternating drug) for the treatment of advanced AIDS. AIDS Clinical Trial Group 193A Study Team. *Journal of Acquired Immune Deficiency Syndromes* **19**, 339-349.

126

- Liang, K.-Y., and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.
- Little, R. J., D'Agostino, R., Cohen, M. L., *et al.* (2012). The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine* **367**, 1355-1360.
- Little, R. J., and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*, 2nd edition. New York: Wiley.
- Mallinckrodt, C. H., Clark, W. S., and David, S. R. (2001). Accounting for dropout bias using mixed-effects models. *Journal of Biopharmaceutical Statistics* **11**, 9-21.
- McCulloch, C. E., Searle, S. R., and Neuhaus, J. M. (2008). *Generalized, Linear, and Mixed Models*. New York: Wiley.
- National Research Council. (2010). *The Prevention and Treatment of Missing Data in Clinical Trials*. Washington, D.C.: National Academies Press.
- O'Neill, R. T., and Temple, R. (2012). The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. *Clinical Pharmacology & Therapeutics* **91**, 550-554.

127

- Rezvan, H. P., Lee, K. J., and Simpson, J. A. (2015). The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. *BMC Medical Research Methodology* **15**, 30.
- Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression-coefficients when some regressors are not always observed. *Journal of the American Statistical Association* **89**, 846-866.
- Royston, P., and White, I. R. (2011). Multiple Imputation by Chained Equations (MICE): Implementation in Stata. *Journal of Statistical Software* **45**, Issue 4.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley.
- Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall.
- van Buuren, S., and Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* **45**, 1-67.
- Verbeke, G., and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer.
- White, I. R., Royston, P., and Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* **30**, 377-399.

128

- 日本製薬工業協会医薬品評価委員会データサイエンス部会．(2014)．臨床試験の欠測データの取り扱いに関する最近の展開と今後の課題について—NASレポート，EMAガイドライン，estimandと解析方法の概説．http://jpma.cstokyo.com/information/evaluation/allotment/missing_data.html
- 日本製薬工業協会医薬品評価委員会データサイエンス部会．(2016)．欠測のある連続量経時データに対する統計手法について（Ver. 2.0）．http://www.jpma.or.jp/medicine/shinyaku/tiken/allotment/pdf/statistics01.pdf
- 高井啓二，星野崇宏，野間久史．(2016). 欠測データの統計科学：医学と社会科学への応用．岩波書店：東京.

129