

高橋将宜・渡辺美智子 著
『欠測データ処理 — Rによる単一代入法と多重代入法 — 』

（共立出版，東京，2017年）

坂田幸繁*

一般的に，経済をはじめ社会科学領域の統計的研究の多くは調査・観測データに依拠しており，調査・観測に基づくデータには欠測が生じることが多い。欠測データを適切に処理しなければ，解析結果に偏りが生じるおそれがある。多重代入法は，尤度解析法と並んで最も汎用的な欠測データ解析法であり，本書は，多重代入法の学術的理論と実用の双方をカバーし，和書としては初めての多重代入法に特化した書籍であり，同時に経済データの利用を全面的に想定したほとんど唯一の欠測処理に関する書籍とってよい。

筆頭著者の高橋将宜会員は，高橋・伊藤（2013, 2014）における多重代入法に関する基礎的研究を出発点に，一貫して，本誌掲載論文を含め，多重代入法に焦点を当てた理論的応用的研究の成果を発表している。これらの一連の研究成果は全体として公的統計データの特性を意識しつつ，他の代入法とも比較衡量しながら，欠測値補定のための実用的なアイデアを提供している。本書は渡辺美智子氏との共著として，欠測データ処理に関するこれまでの最新の知見をもとに新たに書き下ろされたものとしてよい。そのなかでもとくに近年の高橋会員の成果（末尾参考文献に掲載）が余すところなく反映されており，そ

のせいかな本書は教育用図書という外形をとっているものの，内実はそのような研究成果を平易で明解な実務レベルに落とし込んだものであり，学術的な視点を失っておらず，また十分論争的な指摘も垣間見られる。

欠測値問題と本学会とのかかわりを振り返ってみると，センサスに代表される大量観察法の技術過程に関わる教科書的言及を除けば，学会化前史である経済統計研究会時代において，「社会科学としての統計学」の原点のひとつを形成する標本調査論争へのコミットメントがある（例えば岩崎（2018）の該当章など参照のこと）。社会現象への確率的数理の適用の困難性を指摘する他方で，欠測値問題につらなる調査拒否や無回答などによる非標本誤差の存在が確率的な抽出枠の崩れとなり，標本調査データやその利用に対して消極的・否定的立場をとる論拠ともなった。しかし論争自体は大屋（1964）などによる標本調査＝技術論（抽出集計の論理・社会適合性）への包摂とともにほぼ収斂し，しばらくはその文脈で欠測値問題が俎上に上ることはほとんどなかったとってよい。

その後の高度経済成長とともにプライバシー意識の高まりを背景に統計調査環境の悪化問題が顕在化していくなかで，九州大学統計学研究室による統計環境実態調査が企画実施され，そこでは悪化（無回答や調査非協力，不在など）が資本主義的市民社会の発展とと

* 正会員，中央大学経済学部
e-mail : ysakata@tamacc.chuo-u.ac.jp

もに制度化、構造化していく形態や不可避性を明らかにしたものの、その解消の根本的処方箋や統計作成・利用上の方法的な対応可能性を論じることはなかった(九州大学統計学研究室(1979)など)。ただ、統計調査環境研究の延伸として、山田茂会員による調査不詳の研究(山田(2012)など)は注目されてよい。それは、本学会においても国勢調査を含む公的統計における不詳の発生特性に集計値ベースの記述的比較アプローチで迫ったものであり、欠測値処理の視点につながる成果といえる。

いずれにしても本学会の基調となる議論としては、標本調査論争時の早くから欠測値問題を指摘しながら、その後の展開では不在や拒否といった調査非協力が代表される調査環境悪化とその帰結(統計の非反映性)に関心が集まり、公的統計におけるワン・ナンバー・センサス化への議論、極論では調査ベース統計からの脱却(たとえばレジスターベースの統計作成)といった大所的制度論議(浜砂(2009)など)が連なる。議論の方向性はともかく結果的には、集計値としての統計表に含まれる欠測について一定の検討はなされてきたが、主に調査票情報(個票)における欠測値の存在と統計利用における課題についてはほとんど指摘されてこなかった。

そのため欠測値を不可避として内包する、現にある調査票ベースの統計情報(集合)をどのように活用するのか、それは統計作成者の視点でも、利用者の視点でも喫緊の課題であるにもかかわらず、本学会では十分な議論を展開できないままにきた。特に調査票情報や匿名化マイクロデータによる2次利用が統計的研究の有力手段となっている現代において、欠測値に対してどのような態度で向き合うのか、それは分析の成否を規定するファクターである(坂田(2006))。欠測値処理は避けて通れない課題であり、多重代入法に関する本書の成果はそれへの有力な解法の提案であり、

肯定、否定にかかわらず、「社会科学としての統計学」という理論と応用の視点で議論すべき良質な材料を与えるものといえる。

本書の特徴と概要を示すことにしよう。まず、数理展開に軸足を置かず、国別経済マクロデータから構成される実データセットを素材にわかりやすい説明を試みている。データ要素が国別マクロであることは説明の明快さという便宜だけであり、ミクロであろうがマクロであろうが実は差異はない。むしろ、先述のように経済データを素材にしたことが、経済統計を扱う機会が多いものには、前処理を含む多重代入の手順を具体的に提示されているだけに貴重な啓蒙書となっている。特に、右に裾の長い分布に関する欠測データ処理が豊富に扱われており、これは英語の文献を含めて稀有な範例情報といえる。経済学の領域で欠測データを用いた統計分析を行うものには本書の貢献するところは大きい。

また、学術的な貢献だけではなく実際的な効用に重きを置いた展開や説明にも特長を認めざるを得ない。欠測値処理のこれまでの書籍では、内容が数理的な理論のみであったり、具体的な応用例が自然科学の分野に限られていたりするなど、社会科学の実証分析に携わる研究者や実務者が多重代入法を実際に活用するにはハードルが高かったが、本書は、欠測データ解析の理論や数式を扱いながらも、実際的な場面での使用にも耐えるよう、配慮が凝らされている。また、公的統計調査において重要となる単一代入法も扱っており、多重代入法から逆に単一代入法を捉えなおすという意味からも統計実務に関する側面での貢献も大きい。

全体的に本書は、欠測データの扱いについて、その初歩から応用、最新の情報まで、非常に丁寧に書かれた良書である。全章にわたって、手順とともに記述の具体性に腐心されており、初心者のみならず、多重代入法の

理論を理解しやすく読ませる非常に工夫された著書といえる。若干挑発的なQ & Aを中心に据えた「まえがき」と「おわりに」との間につぎの14章に分けて本題が構成される。それらの概要と特記点を以下に示して各論的評価に代える。

- 第1章 Rによるデータ解析
- 第2章 不完全データの統計解析
- 第3章 単一代入法
- 第4章 多重代入法の概要
- 第5章 多重代入法のアルゴリズム
- 第6章 多重代入モデルの診断
- 第7章 量的データの多重代入法Ⅰ：平均値の t 検定
- 第8章 量的データの多重代入法Ⅱ：重回帰分析
- 第9章 質的データの多重代入法Ⅰ：ダミー変数のある重回帰分析
- 第10章 質的データの多重代入法Ⅱ：ロジスティック回帰分析
- 第11章 時系列データの多重代入法：ARIMAモデル
- 第12章 パネルデータの多重代入法：固定効果と変量効果
- 第13章 感度分析：NMARの統計解析
- 第14章 事前分布の導入

第1章では、実データを用いた実用的な分析を行うための環境を整えることと、完全データにおけるデータ分析の復習を兼ねる目的で、Rによるデータ解析を扱っている。

第2章では、欠測データ特有の問題点を議論している。つまり、欠測パターンや欠測メカニズムといった事項を扱い、図解やシミュレーションデータを用いて具体的に論じている。特に、2.5節の「MARについての注意点」では、わが国の欠測データ解析の専門家の中で、MARとは「欠測する値に依存しない欠測」と解説されることがあるが、このような解説が誤っていることを示し、MARとNMAR

の違いは種類の違いではなく程度の違いであることを指摘しており、学術的に興味深い内容を含んでいる。

第3章では、確定的回帰代入法、確率的回帰代入法、比率代入法、平均値代入法、ホットデック法といった単一代入法について扱っている。これらは公的統計調査においても多用されているものであり、欠測値処理の学術研究の基礎をなす手法として議論を整理している。高橋(2017)で報告されているとおり、比率代入法は経済データによく用いられており、ホットデック法は世帯データによく用いられている手法と位置づける。また、3.3節において3種類の比率代入法を一般化した議論は、Takahashi et al. (2017)の研究結果が反映されており、和文の文献では唯一の情報源となっている。

第4章では、多重代入法と単一代入法の違いについて、数式・実例・図・表を用いて、理論的かつ具体的に論じている。特に、4.8.1項の「適切な多重代入法」では、高橋・阿部・野呂(2015)およびTakahashi(2017a)の研究結果に基づき、「多重代入は欠測値に1つの値を代入する単一代入を繰り返し行ったもの」という誤った解説を糺し、適切な多重代入法とは単一代入法を複数回実行したのではなく、欠測データの事後予測分布から無作為抽出したパラメータ推定値を用いたシミュレーションであることを指摘している。多重代入法の本質にかかわる説明を含む重要な章となっている。

第5章では、伝統的なマルコフ連鎖モンテカルロ法(MCMC)に基づくデータ拡大法(data augmentation)、代替法として提案された完全条件付き指定(fully conditional specification)、最新のEMB(expectation-maximization with bootstrapping)アルゴリズムを導入している。特に、最新のEMBアルゴリズムを扱っている和書は少なく、しかもかなり詳細で正確な解説にスペースが割かれており、き

わめて貴重な情報である。さらにこの章では、Takahashi (2017a) に基づき、3つのアルゴリズムの特性を「ジョイントモデリングと条件付きモデリング」、「MCMCと非MCMC」という視点から整理し、3つのアルゴリズムそれぞれの長所と短所を示し、性能比較を行っている。このような視点からの比較は、本書が唯一の事例（下表）である。

	ジョイント モデリング	条件付き モデリング
MCMC	データ拡大法	完全条件付き指定
非MCMC	EMB	

第6章では、上書き代入法、欠測地図、密度の比較といった多重代入モデルの診断方法を扱っている。特に、多重代入モデルの診断方法を3つのアルゴリズムすべてにおいて、統一的に扱った書籍は、和文のみならず欧文においても存在せず、貴重な資料といえる。また、本章後半では、Takahashi (2017b, c), Takahashi et al. (2017) に拠り、対数正規分布データの代入法を扱っている。一般的に、代入モデルは多変量正規分布を仮定することが多いが、対数正規分布データに対して対数変換を施して正規分布で近似した場合、指数変換によって元のスケールに戻すことができないことを指摘し、解決策を提示している。

参考文献

- 岩崎俊夫 (2018), 『社会統計学の伝統と継承：論点と関連論文 (1955-90)』, お茶の水書房。
 大屋祐雪 (1964), 「標本調査の論理」, 『統計学』第12号, pp.15-26。
 九州大学統計学研究室 (1979) 「特集『統計環境実態調査』報告 I」, 『研究所報』(法政大学日本統計研究所)No. 4。
 坂田幸繁 (2006), 「個票データと統計利用」, 『統計学』第90号, pp.31-42。
 高橋将宜 (2017), 「諸外国の公的統計における欠測値の対処法：集計値ベースと公開型マイクロデータの代入法」, 『統計学』第112号, pp.65-83。
 高橋将宜・阿部穂日・野呂竜夫 (2015), 「公的統計における欠測値補定の研究：多重代入法と単一代入法」, 『製表技術参考資料』no. 30, pp.1-95。
 高橋将宜・伊藤孝之 (2013), 「経済調査における売上高の欠測値補定方法について：多重代入法による精度の評価」, 『統計研究彙報』第70号, no. 2, pp.19-86。
 高橋将宜・伊藤孝之 (2014), 「様々な多重代入法アルゴリズムの比較：大規模経済系データを用いた

第7章から第12章は、政治・経済の実データを用いた分析結果とその再現方法を扱っており、計量経済学・計量政治学・データサイエンスといった実証研究分野への多重代入法の普及に貢献している。特に、経済発展を被説明変数として、重回帰モデル、共分散分析モデル、ロジスティック回帰モデル、ARIMAモデル、固定効果モデルと変量効果モデルなど、さまざまな種類のデータ形式に対応するモデリングを扱っている。諸外国の文献を含めて、これほど多重代入法について具体的な内容を扱ったものを評者は知らず、本書の展開は実際のデータ解析に有益な情報を提供している。

第13章では欠測メカニズムが無作為ではないNMARの場合の対応策として、感度分析を扱っている。統計環境Rにおける具体的な実行方法も示しており、極めて実践的であり、感度分析の普及に貢献する内容となっている。第14章では、ベイイズ理論に基づく多重代入法の真価を発揮させるべく、事前分布の導入を扱っている。通常、無難に無情報事前分布を想定する著書は多いが、利用可能な知識を実質的な事前分布に表現して解析に供しようとする本章のような記述はまれであり、ここでの(Rによる)手法と手順の具体的な紹介は実用的で貴重である。

- 分析」,『統計研究彙報』第71号, no. 3, pp.39-82.
- 濱砂敬郎 (2009), 「現代センサス革命の一断面: ドイツの2011年統計登録簿型人口センサスについて」,『熊本学園大学経済論集』15(3), pp.17-34.
- 山田茂 (2012), 「2009年『経済センサス-基礎調査』結果の精度について」,『国士舘大学政経論叢』24(1), pp.1-33.
- Takahashi, M., Iwasaki, M., and Tsubaki, H. (2017), “Imputing the Mean of a Heteroskedastic Log-Normal Missing Variable : A Unified Approach to Ratio Imputation”, *Statistical Journal of the IAOS*, vol. 33, no. 3, pp.763-776.
- Takahashi, M. (2017a), “Statistical Inference in Missing Data by MCMC and Non-MCMC Multiple Imputation Algorithms : Assessing the Effects of Between-Imputation Iterations”, *Data Science Journal*, vol. 16, no. 37, pp.1-17.
- Takahashi, M. (2017b), “Multiple Ratio Imputation by the EMB Algorithm : Theory and Simulation”, *Journal of Modern Applied Statistical Methods*, vol. 16, no. 1, pp.630-656.
- Takahashi, M. (2017c), “Implementing Multiple Ratio Imputation by the EMB Algorithm (R)”, *Journal of Modern Applied Statistical Methods*, vol. 16, no. 1, pp.657-673.

機関誌『統計学』投稿規程

経済統計学会（以下、本会）会則第3条に定める事業として、『統計学』（電子媒体を含む。以下、本誌）は原則として年に2回（9月、3月）発行される。本誌の編集は「経済統計学会編集委員会規程」（以下、委員会規程）にもとづき、編集委員会が行う。投稿は一般投稿と編集委員会による執筆依頼によるものとし、いずれの場合も原則として、本投稿規程にしたがって処理される。

1. 総則

1-1 投稿者

会員（資格停止会員を除く）は本誌に投稿することができる。

1-2 非会員の投稿

- (1) 原稿が複数の執筆者による場合、筆頭執筆者は本会会員でなければならない。
- (2) 常任理事会と協議の上、編集委員会は非会員に投稿を依頼することができる。
- (3) 本誌に投稿する非会員は、本投稿規程に同意したものとみなす。

1-3 未発表

投稿は未発表ないし他に公表予定のない原稿に限る。

1-4 投稿の採否

投稿の採否は、審査の結果にもとづき、編集委員会が決定する。その際、編集委員会は原稿の訂正を求めることがある。

1-5 執筆要綱

原稿作成には本会執筆要綱にしたがう。

2. 記事の分類

2-1 研究論文

以下のいずれかに該当するもの。

- (a) 統計およびそれに関連した分野において、新知見を含む会員の独創的な研究成果をまとめたもの。
- (b) 学術的な新規性を有し、今後の研究の発展可能性を期待できるもので、速やかな成果の公表を目的とするもの。

2-2 報告論文

研究論文に準じる内容で、研究成果の速やかな報告をとくに目的とする。

2-3 書評

統計関連図書や会員の著書などの紹介・批評。

2-4 資料

各種統計の紹介・解題や会員が行った調査や統計についての記録など。

2-5 フォーラム

本会の運営方法や統計、統計学の諸問題にたいする意見・批判・反論など。

2-6 海外統計事情

諸外国の統計や学会などについての報告。

2-7 その他

全国研究大会・会員総会記事、支部だより、その他本会の目的を達成するために有益と

思われる記事。

3. 原稿の提出

3-1 投稿

原稿の投稿は常時受け付ける。

3-2 原稿の送付

原則として、原稿は執筆者情報を匿名化したPDFファイルを電子メールに添付して編集委員長へ送付する。なお、ファイルは『統計学』の印刷レイアウトに準じたPDFファイルであることが望ましい。

3-3 原稿の返却

投稿された原稿（電子媒体を含む）は、一切返却しない。

3-4 校正

著者校正は初校のみとし、大幅な変更は認めない。初校は速やかに校正し期限までに返送するものとする。

3-5 投稿などにかかわる費用

- (1) 投稿料は徴収しない。
- (2) 掲載原稿の全部もしくは一部について電子媒体が提出されない場合、編集委員会は製版にかかる経費を執筆者（複数の場合には筆頭執筆者）に請求することができる。
- (3) 別刷は、研究論文、報告論文については30部までを無料とし、それ以外は実費を徴収する。
- (4) 3-4項にもかかわらず、原稿に大幅な変更が加えられた場合、編集委員会は掲載の留保または実費の徴収などを行うことがある。
- (5) 非会員を共同執筆者とする投稿原稿が掲載された場合、その投稿が編集委員会の依頼によるときを除いて、当該非会員は年会費の半額を掲載料として、本会に納入しなければならない。

3-6 掲載証明

掲載が決定した原稿の「受理証明書」は学会長が交付する。

4. 著作権

4-1 本誌の著作権は本会に帰属する。

4-2 本誌に掲載された記事の発行時に会員であった執筆者もしくはその遺族がその単著記事を転載するときには、出所を明示するものとする。また、その共同執筆記事の転載を希望する場合には、他の執筆者もしくはその遺族の同意を得て、所定の書面によって本会に申し出なければならない。

4-3 前項の規定にもかかわらず、共同執筆者もしくはその遺族が所在不明のため、もしくは正当な理由によりその同意を得られない場合には、本会が承認するものとする。

4-4 執筆者もしくはその遺族以外の者が転載を希望する場合には、所定の書面によって本会に願い出て、承認を得なければならない。

4-5 4-4項にもとづく転載にあたって、本会は転載料を徴収することができる。

4-6 会員あるいは本誌に掲載された記事の発行時に会員であった執筆者が記事をウェブ転載するときには、所定の書類によって本会に申し出なければならない。なお、執筆者が所属する機関によるウェブ転載申請については、本人の転載同意書を添付するものとする。

- 4-7 会員以外の者，機関等によるウェブ転載申請については，前号を準用するものとする。
- 4-8 転載を希望する記事の発行時に，その執筆者が非会員の場合には，4-4，4-5項を準用する。
1997年7月27日制定（2001年9月18日，2004年9月12日，2006年9月16日，2007年9月15日，2009年9月5日，2012年9月13日，2016年9月12日一部改正）

編集委員会からのお知らせ
機関誌『統計学』の編集・発行について

編集委員会

I. 正誤表

本誌第115号(2018年9月発行)において表記に誤りがありましたので、お詫びして訂正します。

- 表紙 (誤) 高部 勲
(正) 高部 勲・山下 智志
裏表紙 (誤) Isao TAKABE
(正) Isao TAKABE, Satoshi YAMASHITA

II. 機関誌『統計学』への投稿を募集しています。

1. 原稿は編集委員長宛に送付して下さい(下記メールアドレス)。
2. 投稿は、常時、受け付けています。なお、書評、資料および海外統計事情等の分類の記事については念のため事前に編集委員長に照会して下さい。
3. 次号以降の発行予定日は次のとおりです。
第117号：2019年9月30日、第118号：2020年3月31日
4. 原則として、すべての投稿原稿が査読の対象となります。投稿に際しては、「投稿規程」および「執筆要綱」の熟読をお願いします。最新版は、本学会の公式ウェブサイトを参照して下さい。
5. 投稿から掲載が決まるまでに要する期間は、通常3ヶ月以上です。投稿にあたっては十分に留意して下さい。
6. 投稿、編集委員会、投稿応募についての問い合わせその他とも、下記編集委員長のメールアドレス宛に送付して下さい。

editorial@jsest.jp

次号以降(2019年度)の編集委員は、つぎのとおりです。

編集委員長 池田 伸(立命館大学)
副委員長 小林良行(総務省統計研究研修所)
編集委員 松川太一郎(鹿児島大学)
水野谷武志(北海学園大学)
山田 満(東北・関東支部)

以上

編集後記

本誌に投稿していただきました執筆者の皆様、そして快く査読をお引き受けいただきました査読者の皆様に改めてお礼申し上げます。上記に示しましたとおり、2019年度から池田編集委員長のもとで、117号と118号が発行されます。引き続き、会員の皆様からの積極的な投稿をお待ちしております。(水野谷武志 記)

執筆者紹介

泉 弘志	(大阪経済大学)	戴 艶娟	(広東外語外貿大学国際経済貿易学院)
李 潔	(埼玉大学経済学部)	平井太規	(神戸学院大学現代社会学部)
高橋雅夫	(独立行政法人統計センター)	坂田幸繁	(中央大学経済学部)

支 部 名

事 務 局

北 海 道	062-8605	札幌市豊平区旭町 4-1-40 北海学園大学経済学部 (011-841-1161)	水野谷武志
東 北・関 東	192-0393	八王子市東中野 742-1 中央大学経済学部 (042-674-3406)	伊藤伸介
関 西	640-8510	和歌山市栄谷 930 和歌山大学観光学部 (073-457-8557)	大井達雄
九 州	870-1192	大分市大字且野原 700 大分大学経済学部 (097-554-7706)	西村善博

『統計学』編集委員

水野谷武志 (北海道) [委員長]	池田 伸 (関西) [副委員長]
小林良行 (東北・関東)	松川太一郎 (九州)
山田 満 (東北・関東)	

統 計 学 No.116

2019年3月31日 発行	発行所	経 済 統 計 学 会 〒112-0013 東京都文京区音羽1-6-9 音羽リスマチック株式会社 TEL/FAX 03 (3945) 3227 E-mail: office@jsest.jp http://www.jsest.jp/
	発行人	代表者 金子治平
	発売所	音羽リスマチック株式会社 〒112-0013 東京都文京区音羽1-6-9 TEL/FAX 03 (3945) 3227 E-mail: otorisu@jupiter.ocn.ne.jp 代表者 遠藤 誠

STATISTICS

No. 116

March 2019

Articles

- International Comparison of Productivity Level by Industry using International
Input-Output Tables
..... Hiroshi IZUMI, Yanjuan DAI and Jie LI (1)

Short Articles

- The Rate and Factors of Husband's Housework in Double-Income Households in Japan
..... Taiki HIRAI (13)

Materials

- The Quality Assurance of Official Statistics in Japan : Framework and Practice
..... Masao TAKAHASHI (26)

Book Reviews

- Masayoshi TAKAHASHI and Michiko WATANABE, *Missing Data Analysis :
Single Imputation and Multiple Imputation in R*, Kyoritsu Shuppan, Tokyo, 2017
..... Yukishige SAKATA (39)

JSES Activities

- JSES Statement on Statistics Act Violations by the Ministry of Health, Labour and Welfare,
Japan (44)
Activities within JSES Branches (46)
Prospects for the Contribution to *Statistics* (51)

JAPAN SOCIETY OF ECONOMIC STATISTICS
