

氏名（本籍） よし もと たく や 吉 本 拓 矢（神奈川県）  
学位の種類 博士（理学）  
学位記番号 甲第 1216 号  
学位授与の日付 2020 年 3 月 17 日  
学位授与の要件 学位規則第 4 条第 1 項該当  
学位論文題目 **Modeling and Decompositions of Symmetry  
for Multi-Way Contingency Tables**  
(多元分割表における対称性のモデリングと  
分解)

論文審査委員 (主査) 教授 富澤 貞男  
教授 宮本 暢子 教授 田中真紀子  
教授 瀬尾 隆 准教授 田畑 耕治

## 論文内容の要旨

統計学における分割表解析は、これまでに多くの論文で統計解析手法の開発に関する議論がなされ、医学、薬学、社会科学など幅広い分野で用いられている。  $T$  種類の分類が同じ  $r$  個の階級から構成されている場合、観測値を  $r^T$  個に分類して得られる表を  $r^T$  分割表という。これは同じ分類からなる多元分割表である。特に、 $T=2$  とした  $r \times r$  の表を正方分割表という。

同じ分類からなる正方分割表や多元分割表の解析では、多くの場合に分類間の相互関連性が強く、統計的独立性は成り立たない傾向にある。そこで、統計的独立性に代わって分類間の対称性に関心があり、対称性の解析が統計モデルを用いて行われる。多元分割表では、1990 年に Bhapkar and Darroch により対称モデル、 $h$  次準対称モデル及び  $h$  次周辺対称モデルが導入された。また、 $h$  次周辺分布 ( $h=1, \dots, T-1$ ) に対して、1990 年に Bhapkar and Darroch により「任意の  $h$  ( $h=1, \dots, T-1$ ) に対して、対称モデルが成り立つための必要十分条件は、 $h$  次準対称モデル及び  $h$  次周辺対称モデルの両方が成り立つことである」という定理が与えられた。また、1 次周辺分布に対して、2002 年に Agresti により 1 次周辺累積ロジスティックモデルが導入され、2007 年に Tahata, Katakura and Tomizawa によって、「1 次周辺対称モデルが成り立つための必要十分条件は、1 次周辺累積ロジスティックモデルと平均一致モデルの両方が成り立つことである」という定理が与えられた。他にも、1976 年に Wall and Lienert により点対称モデルが導入され、2008 年に Tahata and Tomizawa により  $h$  次準点対称モデル及び  $h$  次周辺点対称モデルが導入された。さらに、2008 年に Tahata and Tomizawa に

より、「任意の  $h$  ( $h=1, \dots, T-1$ ) に対して、点対称モデルが成り立つための必要十分条件は、 $h$  次準点対称モデルと  $h$  次周辺点対称モデルの両方が成り立つことである」という定理が与えられ、そしてモデルの直交性に関する定理が与えられた。多元分割表では、対称性及び同等性の確率構造が複雑になるため、それらの確率構造を示す新たなモデルの提案や、モデルの分解に関する新たな定理が必要となる。

本論文は 5 つの章から構成されている。以下に各章の概要を記す。

第 1 章では、多元分割表の研究に関する歴史的背景、並びに本論文の要約を記した。

第 2 章では、多元分割表に対して  $h$  次周辺累積ロジスティックモデルを提案した。これは、Agresti の 1 次周辺累積ロジスティックモデルの拡張である。また、「任意の  $h$  ( $h=1, \dots, T-1$ ) に対して、 $h$  次周辺対称モデルが成り立つための必要十分条件は、 $h$  次周辺累積ロジスティックモデル、 $h-1$  次周辺対称モデル及び  $h$  次モーメント一致モデルのすべてが成り立つことである」という定理を与えた。さらに、「 $T-1$  次周辺対称モデルが成り立つための必要十分条件は、すべての  $h$  ( $h=1, \dots, T-1$ ) に対して  $h$  次周辺累積ロジスティックモデル、及び  $h$  次モーメント一致モデルが成り立つことである」という系を与えた。本章で与えたモデルの分解に関する定理は、2007 年の Tahata et al. の定理の一般化である。提案したモデルと定理は、実データ解析において  $h$  次周辺対称モデルの適合度が悪い場合に、どのような確率構造が存在しているのかを推測するのに有用である。さらに、1984 年と 2016 年の米国の政府支出に関する意見のデータに対して、提案したモデル及びモデルの分解に関する定理を適用し、これまでの方法では得られなかった有用な結果と解釈が得られることを示した。

第 3 章では、多元分割表に対して  $f$  ダイバージェンスに基づく  $h$  次準点対称モデルを提案した。提案したモデルは、一定の条件を与えた下で、点対称性を満たす離散確率分布に対する  $f$  ダイバージェンスを最小にすることで導かれる。また、「任意の  $h$  ( $h=1, \dots, T-1$ ) に対して、点対称モデルが成り立つための必要十分条件は、 $f$  ダイバージェンスに基づく  $h$  次準点対称モデルと  $h$  次周辺点対称モデルの両方が成り立つことである」という定理を与えた。さらに、本章で与えたモデルの分解に関して、「任意の  $h$  ( $h=1, \dots, T-1$ ) に対して、点対称モデルの適合度を検定するための尤度比カイ二乗統計量は、点対称モデルの下で  $f$  ダイバージェンスに基づく  $h$  次準点対称モデルと  $h$  次周辺点対称モデルに対する各尤度比カイ二乗統計量の和に漸近的に同等である」というモデルの直交性に関する定理を与えた。提案したモデル、モデルの分解に関する定理、並びにモデルの直交性に関する定理は、2008 年の Tahata and Tomizawa のモデル及び定理の一般化である。さらに、多元分割表の適用例として、2001 年、2002 年及び 2003 年のある日本の都市における気象データに対して、提案モデル及び定理を適用し、モデルと定理の有用性を示した。

第 4 章では、一つだけが 1 をとり、残りはすべて 0 をとる  $r$  変量の確率変数ベクトルを定義する。この確率変数ベクトルに関する同時確率が、多元分割表のセル確率によって与えられるものとする。このとき、多元分割表に対して  $h$  次モーメント対称モデルを提案した。提案したモデルは、確率変数ベクトルに関する平均まわりの  $h$  次 ( $h=2, \dots, T$ ) モーメントの対称性及び同等性を表すモデルである。また、「任意の  $h$  ( $h=2, \dots, T$ ) に対して、 $h$  次周辺対称モデルが成り立つための必要十分条件は、 $h$  次モーメント対称モデルと  $h-1$  次周辺対称モ

デルの両方が成り立つことである」という定理を与えた。さらに、「対称モデルが成り立つための必要十分条件は、すべての  $h$  ( $h=2, \dots, T$ ) に対して  $h$  次モーメント対称モデルが成り立ち、かつ 1 次周辺対称モデルが成り立つことである」という系を与えた。モデルの分解に関する定理は、実データ解析において対称モデルまたは  $h$  次周辺対称モデルの適合度が悪い場合に、どのような確率構造が存在しているのかを推測するのに有用である。さらに、多元分割表の適用例として、2003 年の日本の 3 都市の気象データに対して、提案したモデル及びモデルの分解に関する定理を適用し、モデルと定理の有用性を示した。

第 5 章では、本論文の結論を述べた。多元分割表において、提案したモデルとモデルの分解に関する定理の有用性を論じた。

## 論文審査の結果の要旨

統計学における分割表解析は、医学、薬学、理学、工学、教育学、心理学、社会学など多くの分野で利用されている。2 種類の分類が同じ分類からなる場合、たとえば、人間の左右裸眼視力を、悪い、やや悪い、やや良い、良い、の 4 つのカテゴリに分けると  $4 \times 4$  分割表を得ることができる。このような 2 種類の同じ分類からなる分割表を正方分割表と呼んでいる。さらに、多種類の分類からなる場合は、多元分割表と呼んでいる。

分割表解析において、我々は全体（母集団）の一部として得られた標本に基づく観測度数しか得ることができない。我々の関心は、母集団における未知の確率分布がどのような構造になっているのかを（母集団の一部として）得られた観測度数から、できる限りわかりやすい解釈が得られるように、高い信頼度で推測することにある。そのためには、データに良く適合し、かつ解釈が容易な確率分布に関する統計モデル（仮説）を導入する必要がある。さらに、モデルがデータに良く適合するかどうかを調べるための検定統計量の開発、モデルの未知パラメータの推定法、モデルの分解など多くの解決すべき問題点がある。

分割表の解析において、多くの人の関心の一つは、分類間が独立（関連性がない）かどうかである。しかし、同じ分類からなる正方分割表や多元分割表の解析においては、分類間の関連性は極めて強く、統計的独立性は成り立たない。代わって、分類間の対称性に関心がある。そのために、対称性あるいは非対称性に関する統計モデルを用いた解析が行われる。歴史的には、Bowker (1948) が対称モデルを導入したのが始まりである。さらに対称モデルを拡張した種々のモデルが提案されている。たとえば、Stuart (1955) の周辺同等モデル、Caussinus (1965) の準対称モデル、Wall and Lienert (1976) の点対称モデルなどがある。また、Caussinus (1965) は正方分割表において「対称モデルが成り立つための必要十分条件は、準対称モデルと周辺同等モデルの両方が成り立つことである」という対称モデルの分解定理を与えた。Bhapkar and Darroch (1990) は多元分割表へこの定理を拡張した。多元分割表において、Tahata et al. (2007) は 1 次周辺対称モ



デルの分解に関する定理を, Tahata and Tomizawa (2008)は点対称モデルの分解に関する定理を与えた.

本論文は, 多元分割表において, 対称性に関するモデルの提案, モデルの分解定理, そしてモデルの適合度検定統計量の直交性に関する定理を与えている. 本論文は, 5つの章から構成されている.

第1章は, 本研究に関する歴史的な背景と各章の概要について述べている.

第2章では, 多元 ( $T$  元) 分割表において, 任意の  $h$  ( $h=1, \dots, T-1$ ) に対して,  $h$  次周辺累積ロジスティックモデルを提案し, 「 $h$  次周辺対称モデルが成り立つための必要十分条件は,  $h$  次周辺累積ロジスティックモデル,  $h-1$  次周辺対称モデル及び  $h$  次モーメント一致モデルのすべてが成り立つことである」という定理を与えている. この定理は Tahata et al. (2007)の定理の一般化である. さらに実際の応用例と共に提案されたモデルと定理の有用性が示されている.

第3章では, 多元分割表において,  $f$ -ダイバージェンスに基づく  $h$  次準点対称モデルを提案している. さらに「点対称モデルが成り立つための必要十分条件は,  $f$ -ダイバージェンスに基づく  $h$  次準点対称モデルと  $h$  次周辺点対称モデルの両方が成り立つことである」という定理を与えている. これらは Tahata and Tomizawa (2008)のモデルと定理の一般化である. そしてモデル間の適合度検定統計量の直交性も成り立つことを証明している. さらに実際の応用例と共に提案されたモデルと分解定理の有用性が示されている.

第4章では, 多元分割表において,  $h$  次モーメント対称モデルを提案し, 「 $h$  次周辺対称モデルが成り立つための必要十分条件は,  $h$  次モーメント対称モデルと  $h-1$  次周辺対称モデルの両方が成り立つことである」という定理を与えている. この定理を用いると, 実データにおいて, 従来の解析法では得られない詳細な解析をすることが可能となり, 得られた定理は極めて有用である.

第5章では, 本論文の結論が述べられている.

以上, 本論文は理論面と応用面の両面において大変高く評価できるものであり, 分割表統計解析の分野に, 独創的な新しい解析法を与えており, この分野に大きな貢献をしている. よって理学的に価値ある知見と成果を得たもので博士 (理学) の学位論文として十分価値あるものと認める.