氏名（本籍）	倉上弘幸（くらかみひろゆき）（埼玉県）
学位の種類	博士（理学）
学位記番号	甲第1052号
学位授与の日付	平成26年3月20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	**On Decompositions of Symmetry for Ordinal Contingency Tables** **（順序分割表における対称性の分解について）**
論文審査委員	（主査）教授　富澤　貞男
	教授　戸川　美郎　　教授　明石　重男
	准教授　宮本　暢子　　教授　山崎多恵子
	教授　瀬尾　　隆

論文内容の要旨

カテゴリカルデータ解析についての多くの論文によって，20 世紀初頭より，2 つのカテゴリ変数間の関連性を要約するのに適切な方法について議論がされてきた．その後，カテゴリカルデータ解析の統計的手法の発展は，社会学や生物医学における調査研究において発生したデータを解析する必要性によって促進されてきた．カテゴリカルデータ解析は，行動科学や疫学，公衆衛生学，教育学，経済学など幅広い分野で用いられている．

カテゴリ変数が 2 つ以上のカテゴリを持つとき，変数を 2 つのタイプに区別する．カテゴリ間に自然な順序を持たない変数を名義変数と呼ぶ．たとえば，血液型（A・B・O・AB）などが該当する．カテゴリ間に自然な順序を持つ変数を順序変数と呼ぶ．たとえば，裸眼視力（良い・やや良い・やや悪い・悪い）などが該当する．この区別はどの統計的手法が適切であるかを決定する．たとえば，名義変数に対する統計的手法は，カテゴリの順序を無視することにより順序変数に対して用いても構わないが，順序変数に対する統計的手法は，順序情報を持っていない名義変数に適用することはできない．

それぞれ r 個のカテゴリと c 個のカテゴリを持つ 2 つの変数を考える．両方の変数における対象の分類の起こりうる結果の組み合わせは r×c 個ある．これを r 行 c 列の長方形型の表にまとめたものを r×c 分割表と呼ぶ．たとえば，左右の裸眼視力を良い・やや良い・やや悪い・悪いと 4 つのカテゴリに分けるとすると，4×4 分割表を得ることができる．また，k 種類の変数がそれぞれ r_i（i=1,…,k）個のカテゴリを持つとき，r_1×・・・r_k 分割表を得

ることができる．これは多元分割表と呼ばれる．

実際の分割表解析において，我々は観測度数しか得ることができない．我々の関心は，その観測度数の背後にある未知の確率分布を高い信頼度で推測することにある．そして，変数間の相互関連性に関して，有用かつわかりやすい解釈を実際のデータに対して与えることが分割表解析の目的である．そのために，データに良く適合し，解釈が容易である統計モデルを導入する必要がある．さらに，モデルにおける未知パラメータの推定法やモデルの適合の良さを計る検定統計量の開発，モデルの分解など解決すべき多くの課題が存在する．本研究は分割表の中でも，同じ分類からなる正方分割表や多元分割表において，これらの課題を取り扱った研究である．

2 元分割表解析において，多くの人の関心は，行変数と列変数が独立かどうか（関連性がないかどうか）にある．しかし，2 種類が同じ分類からなる正方分割表では，多くの場合分割表の主対角セル付近に観測度数が集中するため，分類間の相互関連性は極めて強く，統計的独立性は成り立たないことが多く見られる．そこで統計的独立性に代わって，対称性や非対称性に関する統計モデルが提案され，解析が行われてきた．Bowker（1948）が分割表のセル確率の対称性の構造を示す，対称モデルを導入したことがその始まりのようである．その後，対称モデルを拡張したモデルとして様々なモデルが導入されている．対称性の構造を示すモデルとして，たとえば，Stuart（1955）の周辺同等モデルや，Caussinus（1965）の準対称モデルが導入された．非対称性を示すモデルとしては，McCullagh（1978）の条件付き対称モデル，Goodman（1979）の対角パラメータ対称モデル，Agresti（1983）の線形対角パラメータ対称モデルなどが提案されてきた．また，Caussinus（1965）は「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．モデルの分解は，分割表をより詳細に解析するのに有用である．

本論文は 4 つの章から構成されており，順序カテゴリ正方分割表および多元分割表において，非対称性に関するモデルの提案や分解定理の導出を扱っている．

第 1 章では，正方分割表の研究に関する歴史的背景と各章の概要について述べた．

第 2 章では，分割表の主対角線に対して非対称な構造を示す，一般化指数対称モデルを導入した．このモデルは，既存の非対称性に関するモデルである，条件付き対称モデルや線形対角パラメータ対称モデルを含んだ形で一般化されている．また，セル確率の重み付きの和の対称性に関するモデルも導入した．このモデルは，Read（1977）のグローバル対称モデルなどを含む形で一般化されたモデルである．この 2 つの導入したモデルを用いて，対称モデルの分解定理を与えた．さらに，「対称モデルの適合度検定統計量が，分解した 2 つのモデルの適合度検定統計量の和と漸近的に等しい」という適合度検定統計量に関する分解を得た．この分解定理と適合度検定統計量の分解についても，Read（1977）の，条件付き対称モデルとグローバル対称モデルを用いた対称モデルの分解や，Tahata and Tomizawa（2008）の，線形対角パラメータ対称モデルを用いた対称モデルの分解を含む形で一般化されていることを示した．また，一般化指数対称モデルの枠組みに含まれるモデルであれば分解定理と適合度検定統計量の分解が成り立つことを示した．そして，導入したモデルと分解定理から有用な結果と解釈が得られることを，複数の実データ解析ととも

に示した．

第 3 章では，周辺分布関数に関するいくつかのモデルを導入した．順序カテゴリ正方分割表において，周辺同等モデルは，行変数の周辺分布関数と列変数の周辺分布関数が同等であるという構造を表す．実際のデータ解析において，周辺同等モデルが成り立たないとき，我々は行変数の周辺分布関数と列変数の周辺分布関数にどのような違いがあるのかに興味がある．Agresti（1984）は周辺分布関数の対数オッズである周辺累積ロジットを用いて，周辺同等モデルの拡張である周辺累積ロジスティックモデルを導入した．Miyamoto, Niibe and Tomizawa（2005）は分割表の行変数と列変数のそれぞれの平均が等しいという構造を示す，平均一致モデルを導入し，「周辺同等モデルが成り立つための必要十分条件は，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」という分解定理を与えた．第 3 章では周辺累積ロジスティクモデルをさらに拡張したモデルを導入した．このモデルは行変数の周辺累積ロジットと列変数の周辺累積ロジットの差がカテゴリの値の線形関数である構造を示しており，行変数と列変数の周辺分布関数の関係を推測するのに有用であると考えられる．行変数と列変数が同じ値を取らないという条件の下での周辺累積ロジスティックモデルも Miyamoto et al.（2005）によって導入されており，そのモデルに関しても拡張したモデルを導入した．また，新たに導入したモデルに関して，多元分割表への拡張を行った．そして，新たに導入したモデルを用いて，周辺同等モデルの分解定理を与えた．この分解定理により，既存の分解定理では解釈を得られなかったような場合でも，より詳細な解析をすることが可能になった．さらに，導入したモデルと分解定理が周辺分布関数の構造に関する考察を与えるのに有用であるということを，実データを用いて示した．

第 4 章では，多元分割表において，第 3 章で扱った周辺累積ロジスティックモデルをさらに一般的に拡張したモデルを導入した．このモデルは，任意の 2 つの周辺累積ロジットの差がカテゴリの値の多項式関数である構造を示すモデルであり，周辺累積ロジスティックモデルや，第 3 章で導入した拡張周辺累積ロジスティックモデルを含む形で一般化されている．このモデルは各変数の周辺分布関数の関係をより詳細に推測するのに有用であると考えられる．また，各変数が同じ値を取らないという条件の下で，周辺累積ロジスティックモデルを一般的に拡張したモデルを導入した．そして，各変数のモーメントの一致に関するいくつかのモデルを導入し，周辺同等モデルの分解定理を複数与えた．ここで与えた分解定理は，Tahata, Katakura and Tomizawa（2007）が与えた周辺累積ロジスティックモデルを用いた分解定理や，第 3 章で与えた拡張周辺累積ロジスティックモデルを用いた分解定理などを含む形での一般化となっている．この分解定理は，実際のデータ解析において周辺同等モデルの当てはまりが悪いとき，その原因をより細かく考察することに有用であると考えられる．さらに，導入したモデル，分解定理の有用性を，理論面だけではなく実データを用いた解析例とともに示した．

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，社会学など幅広い分野で利用されている．ある人間の集団を喫煙習慣の有無と肺疾患の有無によって4つの小集団に分けると，各集団の人数からなる2×2分割表を得ることができる．また，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4×4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）の一部分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（母集団の一部分として）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，モデル間の関連性，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（母数）の推定法，モデルの分解，モデルからの隔たりを測る尺度の開発など多くの解決すべき問題点がある．

一般の2元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．たとえば，人間は，幼児の頃は左右裸眼視力は対称的であるが，成長とともに対称性は崩れる傾向にあり，左右どちらが良い目なのか，また，どのように非対称になるのかに関心がある．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．その後，対称モデルを拡張した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) の準対称モデルなどがある. また, Caussinus は「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．

本論文は，分割表解析において，非対称性に関するモデルの提案，モデルの直交分解定理，そして周辺累積ロジスティックモデルの拡張とそれを用いての周辺同等モデルの分解定理を与えている．本論文は，4つの章から構成されている．

第1章では，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，正方分割表において一般化指数対称モデルを提案し，「対称モデルが成り立つための必要十分条件は，一般化指数対称モデルと重み付き周辺対称性のモデルの両方が成り立つことである」という定理を与え，さらに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．実際の応用例と共に直交分解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，拡張周辺累積ロジスティックモデルを提案している．このモデルは任意の2つの周辺累積ロジットの差がカテゴリ値の線形構造を示している．

そして「周辺同等モデルが成り立つための必要十分条件は，拡張周辺累積ロジスティックモデル，平均一致モデルそして分散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．また，観測値が主対角セルに入らないという条件の下で同様な定理が成り立つことも示している．これらの分解定理は，データ解析において周辺同等モデルが成り立たない場合に，より詳細に解析するのに有用であり，実際の応用例と共にそれらが示されている．

第４章は，第３章を一般化したものである．一般化周辺累積ロジスティックモデルを提案している．このモデルは任意の２つの周辺累積ロジットの差がカテゴリ値の多項式の構造を示している．そして「周辺同等モデルが成り立つための必要十分条件は，一般化周辺累積ロジスティックモデル，いくつかの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，実データの解析を行う際，従来の解析法では得られない詳細な解析が可能であり，極めて有用である．

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．