

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 3 日現在

機関番号：32660

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24540145

研究課題名(和文) 多元分割表解析における対称性のモデリングと尺度の分解

研究課題名(英文) Model and measure of symmetry for multiway contingency tables

研究代表者

富澤 貞男 (TOMIZAWA, Sadao)

東京理科大学・理工学部・教授

研究者番号：50188778

交付決定額(研究期間全体)：(直接経費) 1,800,000円

研究成果の概要(和文)：(1) 2元の正方分割表あるいは多元分割表において対称性や非対称性に関するいくつかのモデルを提案した。(2) 対称性や非対称性のモデルを2つ(あるいは3つ以上)のモデルへの分解定理を与えた。(3) 対称モデルの分解において適合度検定統計量の直交分解を与えた。(4) 正方分割表において対称性(あるいは非対称性)のモデルからの隔たりを測る尺度を提案した。(5) 連続型多変量解析として多変量確率密度関数関しての対称な確率密度関数を提案し、対称確率密度関数の分解を与えた。

研究成果の概要(英文)：(1) For two-way square contingency tables or multi-way contingency tables, we have proposed some models of symmetry or asymmetry; (2) we have given decomposition theorem of symmetry (asymmetry) model into two (or three more) models; (3) we have given the orthogonal decomposition of goodness-of-fit test statistic for the symmetry (or asymmetry) model; (4) we have proposed some measures for symmetry (or asymmetry); and (5) for continuous multivariate analysis, we have proposed the symmetry of multivariate probability density function, and given its decomposition.

研究分野：数理統計学

キーワード：分割表統計解析

## 1. 研究開始当初の背景

行分類と列分類が同じ正方分割表において未知のセル確率の対称性や非対称性の構造を述べたモデルは、著者の研究も含めていると提案されている。特に、Bowker(1948)の対称(S)モデル、Causinuss(1965)の準対称(QS)モデル、Stuart(1955)の周辺同等(MH)モデルは、正方分割表解析の分野では良く知られている。また、これらは対称性の構造を述べたモデルである。

一方、非対称性の構造を述べたモデルとして、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やAgresti(1983)の線形パラメータ対称モデルなどがある。

また、モデルの分解としてCausinuss(1965)の「Sモデルが成り立つための必要十分条件は、QSモデルとMHモデルの両方が成り立つことである」という分解定理は良く知られている。正方分割表解析における種々の対称性や非対称性のモデルに関しては、たとえば、富澤(2006,日本数学会「数学」の論説)に詳細に述べられている。これらのモデルは、分割表における主対角線に関する対称性や非対称性のモデルである。一方、Wall and Lienert(1976)は分割表の中心点(あるいは中心セル)に関する点対称性を示すモデルを提案した。Tomizawa(1986)は対称と点対称を同時に見る二重対称モデルを提案した。

一方、3元以上の多元分割表における対称性や非対称性のモデルの研究はあまり進んでいない。上記のCausinuss(1965)の2元(正方)分割表の研究の多元分割表への拡張は、Bhaskar and Darroch(1990)により多少は考えられている。Tahata and Tomizawa(2008)は、多元分割表において点対称(PS)モデル、準点対称(QPS)モデル、周辺点対称(MPS)モデルを新たに導入し、「PSモデルが成り立つための必要十分条件は、QPSモデルとMPSモデルの両方が成り立つことである」というPS

モデルの分解定理を与えた。しかし、多元分割表において、種々の非対称性の構造を述べるモデルや非対称モデルの分解の研究は進んでいない。

データ解析を行うとき、関心あるモデルが成り立たないときは(適合度が悪いとき)、さらに条件を緩めたモデルの当てはめや残差分析などに一般には関心がある。しかし、相関係数のようにモデルからの(関心ある)最大の隔たりの構造を定義し、その最大の隔たりの方向へ向かって、関心あるモデルからどれくらい確率構造がずれているのかを測ることに大変関心がある。2元分割表において対称モデルからの隔たりを測る尺度を導入したのは、世界で著者(Tomizawa,1994)が提案したのがはじめてである。その後、周辺同等モデルからの隔たりを測る尺度も著者(Tomizawa,1995)が提案した。その後、いくつかの対称性に関する尺度を提案した(たとえば富澤(2006,日本数学会「数学」の論説)を参照されたい)。

## 2. 研究の目的

(1)多元分割表解析において、種々の対称性や非対称性を述べるセル確率に関する一般的な対称性のモデルのクラス(GSと記す)を導入し、更に分割表の内部の確率構造に関する一般的な(たとえばオッズ比に基づく)準対称モデル(GQSと記す)と周辺確率に関する一般的周辺対称モデル(GMと記す)のクラスを導入すること、特に「GSモデルが成り立つ必要十分条件は、GQSモデルとGMモデルの両方が同時に成り立つことである」という分解定理を考える。

(2)多元分割表で、一般的対称性、非対称性モデルからの隔たりを測る尺度とその分解を考える。

(3)多元分割表で、GSモデルのGQSモデルとGMモデルへの分解を導入する際、グラフィカルモデリングとの関係を考える。

(4)多元分割表で、幾何学的アプローチが

ら対称性のモデルの直交性を考える。

### 3. 研究の方法

- (1) 多元分割表で種々の対称性モデルを含む一般的な対称性のモデル(クラス)を提案する。
- (2) 多元分割表において一般的な対称性モデルに対する分解を与える。
- (3) 多元分割表において対称性や非対称性からの隔たりを測る尺度を提案する。
- (4) グラフィカルモデリングに関する文献講読, 分離可能モデルにおける対称性のモデルの分解を考える。
- (5) 情報幾何学に関する文献講読, 情報幾何学的アプローチからの直交概念を研究する。
- (6) 多次元離散データ収集(または連続データで離散型のデータに直す)・文献調査する。
- (7) 研究成果を論文にまとめ, 学術雑誌へ投稿する
- (8) 学会, シンポジウム等で研究発表する。

### 4. 研究成果

(1) 正方(2元)または多元分割表での種々の対称性(非対称性)に関するモデルを提案した。具体的には, 例えば, Bowker (1948)の対称モデルを拡張したモデルとして McCullagh (1978)のパリンドロミック(palindromic)対称モデルがあるが, 本研究ではそれらのモデルを含む一般化したパリンドロミック対称モデルを提案した。

また, 正方分割表解析において2変量  $t$  分布型対称モデルを提案した。これまで2変量正規分布型対称モデルは著者が過去に提案したが, 本研究ではそれとは異なるモデルである。提案したモデルは, 潜在連続分布として2変量  $t$  分布が考えられる場合, 切断点を設けて分割表を作成し, 離散型分布にすると良く適合するモデルである。これらのモデルの有用性はシミュレーションでも示し, 非常に優れたモデルであることを示した。

また, 正方分割表において, 不完備型対称モデルも提案した。このモデルは, 完備型対称モデル(従来の対称モデル)が成り立たない場合に, それよりも制約の弱いモデルで有り, 主対角線に関していくつかの対称なセルを除いての確率の対称性を示すモデルである。

さらに, 順序カテゴリ正方分割表において拡張した非対称モデルを提案した。具体的には, McCullagh (1978)のパリンドロミック対称モデルを一般の  $m$  個のパラメータをもつ拡張パリンドロミック対称モデルを提案した。そして具体例として, 父親とその息子の学歴データに提案したモデルが良く当てはまることを示し, また提案したモデルの有用性を示した。

一方, 順序カテゴリ正方分割表において修正(拡張でなく)パリンドロミック対称モデルを提案した。拡張パリンドロミック対称モデルは, 対称的な累積確率の対数比率がカテゴリの多項式として表せるのに対して, 修正パリンドロミック対称モデルは, 対称的な累積確率の比率が多項式として表せるというモデルである。これらのモデルが良く適合する例も与えた。

また, 順序カテゴリ正方分割表において Tomizawa (1992)は対角指数対称モデルを提案したが, 本研究ではその拡張である準対角指数対称モデルを提案した。そして, 対角指数対称モデルは準対角指数対称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を与えた。また, モデルの適合度を示す検定統計量の直交性が成り立つことも証明付きで示した。シミュレーションにより提案したモデルやモデル分解の有用性を示した。

正方分割表や多元分割表の解析において, 対称性や非対称性の確率構造を探るモデルはこれまであまりなく, 本研究で提案されたモデルは多元分割表データを解析する上

で今までの解析法では得られなかった詳細な解析ができ、今後これらの研究成果は多くの研究に応用されると期待できる。

(2) 正方分割表において対称性の3つのモデルへの分解を提案し、対称モデルが成り立つための必要十分条件は、累積2比パラメータ対称モデル、グローバル対称モデルと周辺平均一致モデルが成り立つことである、という定理を与えた。また、別の分解として、対称モデルが成り立つために必要十分条件は、パ lindロミック対称モデル、周辺平均一致モデルと累積部分対称モデルが成り立つことである、という定理を与えた。更には、別の分解として、対称モデルが成り立つための必要十分条件は、修正パ lindロミック対称モデル、周辺平均一致モデルと累積対称モデルが成り立つことである、という定理を与えた。

モデルの分解を用いた研究はこれまであまりなく、本研究で得られたモデルの分解に基づく新しいデータ解析法は、従来の解析法では得ることのできない詳細なデータ解析を可能とするものである。

(3) 分割表解析は離散型多変量解析であるが、連続型多変量解析として、多変量密度関数に関して対称的な確率密度関数を提案し、対称確率密度関数の分解を与えた。具体的には準対称密度関数を提案し、対称密度関数が成り立つための必要十分条件は、準対称確率密度関数と周辺対称密度関数の両方が成り立つことである、という定理を与えた。

更に、多変量点対称密度関数、準点対称密度関数、そして周辺点対称密度関数を提案し、点対称密度関数が成り立つための必要十分条件は、準点対称密度関数と周辺点対称密度関数が成り立つことである、という定理を与えた。また、多変量正規分布に対してそれらの関係がどのようになっているのかを示した。

(4) 正方分割表において対称性(非対称性)のモデルが成り立たないとき、そのモデルからの隔たりがどの程度なのかを測る尺度を提案した。周辺の行確率分布と列確率分布の同等性(周辺同等性または周辺対称性)からの隔たりを測る尺度を提案した。

また、Arimoto エントロピーを用いて対称性からの隔たりを測る尺度を提案した。

更に、2元あるいは多元分割表において、予測誤差に関する比例縮小度を測る尺度を与えた。例えば、2元分割表において目的変数Yを予測するのに説明変数Xの値(情報)を知らなかったときよりも知ったときの方が、予測誤差がどれくらい縮小するかを測る尺度を順序カテゴリーの場合に提案した。

尺度は未知なので、提案した尺度の推定法も与え、そして推定した尺度の近似分散を理論的に求め、尺度の信頼区間を理論的に導出した。さらに、具体的にいくつかの実データに適用し、提案した尺度の有用性を示した。特に異なるいくつかのデータに対してモデルからの隔たりの程度を比較するのに提案した尺度は極めて有用である。

(5) グラフィカルモデリングに関する文献調査と基礎研究を行い、基礎知識を深め、対称モデルとの関係を考察した。

(6) 情報幾何学に関する文献調査と基礎研究を行い、基礎知識を深め、対称性のモデルの直交分解、つまり、検定統計量の漸近的直交性との関係を考察した。

(7) 研究成果を論文にまとめ海外の査読付き学術雑誌や国内の査読付き学術雑誌へ投稿し、査読結果に基づき修正し、論文として発表した。また、現在も関連ある研究を継続中である。

(8) 研究成果を日本数学会や日本統計学会、応用統計学会、日本計算機統計学会などの統計関連の学会や統計関連の科研費に基づくシンポジウムで発表した。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 15 件)

Yamamoto, K. and Tomizawa, S. (2012): Statistical analysis of case-control data of endometrial cancer based on new asymmetry models. Journal of Biometrics and Biostatistics, Vol. 3, pp.1-4. 査読有 .

<http://omicsonline.org/editorialboard-biometrics-biostatistics-open-access.php>

Tahata, K., Yamamoto, K., and Tomizawa, S. (2012): Decomposition of symmetry using palindromic symmetry models in two-way classification. Journal of Statistics Applications and Probability, Vol.1, pp.175-178. 査読有

<http://www.naturalspublishing.com/show.asp?JorID=3&pgid=0>

Yamamoto, K., and Tomizawa, S. (2012): Analysis of tooth decay data in Japan using asymmetric statistical models. Open Access Medical Statistics, Vol. 2, pp.6. 査読有 .

<http://www.dovepress.com/open-access-medical-statistics-journal>

Iki, K., Tahata, K., and Tomizawa, S. (2012): Decomposition of symmetric multivariate density function. SUT Journal of Mathematics, Vol.48, pp.199-211. 査読有 .

<http://www2.ma.kagu.tus.ac.jp/index.php?SUT%20Journal%20of%20Mathematics>

Iki, K., Ishihara, T., and Tomizawa, S. (2013): Bivariate t-distribution type symmetry model for square contingency tables with ordered categories. Model Assisted Statistics and Applications, Vol.8, pp.315-319. 査読有 .

<http://www.iospress.nl/journal/model-assisted-statistics-and-applications/>

Yamamoto, K., Murakami, K., and Tomizawa, S. (2013): Point-symmetry models and decomposition for collapsed tables. Journal of Applied Statistics. Vol.40, pp.1446-1452. 査読有 .

[http://www.tandfonline.com/loi/cjas20#.VW6077kw\\_cs](http://www.tandfonline.com/loi/cjas20#.VW6077kw_cs)

Saigusa, Y., Tahata, K., and Tomizawa, S. (2014): An extended asymmetry model for square contingency tables with ordered categories. Model Assisted Statistics and Application, Vol. 9,

pp.151-157. 査読有 .

<http://www.iospress.nl/journal/model-assisted-statistics-and-applications/>

Iki, K., and Tomizawa, S. (2014): Point-symmetric multivariate density function and its decomposition. Journal of Probability and Statistics, Vol.2014, pp.1-6. 査読有 .

<http://www.hindawi.com/journals/jps/>

Iki, K., Yamamoto, K., and Tomizawa, S. (2014): Quasi-diagonal exponent symmetry model for square contingency tables with ordered categories. Statistics and Probability Letters, Vol.92, pp.33-38. 査読有 .

<http://www.journals.elsevier.com/statistics-and-probability-letters/>

Tahata, K., and Tomizawa, S. (2014): Symmetry and asymmetry models and decompositions of models for square contingency tables. SUT Journal of Mathematics, Vol.50, pp.131-165. 査読有

<http://www2.ma.kagu.tus.ac.jp/index.php?SUT%20Journal%20of%20Mathematics>

[学会発表](計 15 件)

川崎協, 田畑耕治, 富澤貞男 (2012年5月20日): 正方分割表における条件付き周辺分布を用いた非対称性の尺度. 応用統計学会, 統計数理研究所 (東京都立川市).

田畑耕治, 山本紘司, 富澤貞男 (2012年9月25日): Decomposition of symmetry using palindromic symmetry model for square contingency tables. 日本数学会, 九州大学理学部 (福岡市). 三枝祐輔, 田畑耕治, 富澤貞男 (2013年3月22日): Extended palindromic symmetry models for square contingency tables with ordered categories. 日本数学会, 京都大学理学部 (京都市).

島田文香, 山本紘司, 富澤貞男 (2013年5月24日): 正方分割表における併合した表を用いた対称性に関する尺度. 応用統計学会, 福島市飯坂温泉観光会館.

小田剛士, 生亀清貴, 富澤貞男 (2013年9月8日): 正方分割表における修正パ Lindromic 対称モデル. 統計関連学会, 大阪大学基礎工学部 (大阪府豊中市).

三枝祐輔, 田畑耕治, 富澤貞男 (2013年9月27日): Decomposition of symmetry using extended palindromic symmetry models for square contingency tables. 日本数学会, 愛媛大学理学部 (愛媛県松山市).

中根弘貴, 田畑耕治, 富澤貞男 (2014年5月23日): 順序カテゴリ正方分割表における2変量正規分布型対称モデル. 応用統計学会, 統計数理研究所 (東京都立川市).

三枝祐輔, 田畑耕治, 富澤貞男 (2014年9月10日): 正方分割表におけるfダイバージェンスに基づく順序準対称モデル. 統計関連学会, 東京大学経済学部 (東京都文京区).

生亀清貴, 富澤貞男 (2014年9月25日): 多変量密度関数の点対称性と分解. 日本数学会, 広島大学理学部 (広島県東広島市).

長縄真学, 田畑耕治, 富澤貞男 (2014年9月10日): 順序カテゴリ正方分割表における拡張線形非対称モデル. 日本数学会, 広島大学理学部 (広島県東広島市).

〔図書〕(計 0件)

〔産業財産権〕

出願状況(計 0件)

名称:

発明者:

権利者:

種類:

番号:

出願年月日:

国内外の別:

取得状況(計 0件)

名称: 発明者:

権利者:

種類:

番号:

出願年月日:

取得年月日:

国内外の別:

〔その他〕

ホームページ等

<http://www.rs.noda.tus.ac.jp/stomizaw/>

## 6. 研究組織

### (1) 研究代表者

富澤 貞男 (TOMIZAWA, Sadao)

東京理科大学・理工学部・教授

研究者番号: 50188778

### (2) 研究分担者

( )

研究者番号:

### (3) 連携研究者

( )

研究者番号: