審 査 の 結 果 の 要 旨

論 文 提 出 者 氏 名　　　小 川　 光 紀

統計モデルの推測においては、検定統計量や推定量などの統計量の標本分布を標本サイズが十分大きい場合の大標本理論を用いて近似することがしばしばおこなわれる。しかしながら、データがスパースとなるような状況では、大標本理論による近似の精度が良くない場合がある。特に観測値が離散となる離散モデルにおいては、有限標本の正確な離散分布に基づく推測が望ましい。この問題に関して、1998 年の Diaconis と Sturmfels の論文において、トーリックイデアルのグレブナー基底の理論に基づく新たなアプローチが提案され、それ以来「代数統計学」とよばれる分野が急速に発展してきた。代数統計学の中の研究テーマとして、本論文で扱われている正確検定のためのマルコフ基底、および最尤推定のためのホロノミック勾配法の二つのテーマは、統計の応用上重要なものである。理論的な観点からは、正確検定のためのマルコフ基底は多項式環のトーリックイデアルのグレブナー基底計算、ホロノミック勾配法は確率分布の正規化定数が満たす微分方程式が定める微分作用素環のイデアルのグレブナー基底計算と関係しており、本論文の研究は理論的な観点からも重要である。

本論文は 「Algebraic Statistical Methods for Conditional Inference of Discrete Statistical Models」（離散統計モデルの条件付き推測問題に対する代数統計的手法）と題し、６章からなる。

第１章「Introduction」(序論)では、代数統計学の分野の発展の概観を与えるとともに、本論文で扱われているいくつかの具体的な統計モデルに関する研究の流れ、理論上および応用上の意義についてまとめている。

第２章「Markov bases for typical block effect models of two-way contingency tables」（２元分割表の典型的なブロック効果モデルに対するマルコフ基底）では、二元分割表の統計モデルを扱っている。二元分割表に関しては独立モデルが基本的な重要性を持つモデルであるが、ここでは独立モデルに部分表の効果を追加した統計モデルを考え、そのマルコフ基底の構造を解析している。二元分割表の独立モデルについては，2次の移動（周辺和を不変に保つ整数ベクトル）の集合がマルコフ基底をなすことが基本的事実として知られている。しかしながら、分割表の一部のブロックに交互作用効果を加えると、対応する統計モデルのマルコフ基底の構造は非自明なものとなる。先行研究の Hara et al. (2009) では一つのブロックの効果を加味した場合についての結果を与えているが、本章では、より一般に複数のブロックの効果を加味した場合に対して、実用上自然ないくつかの設定の下でマルコフ基底の具体形の導出に成功している。

第３章「Graver basis for an undirected graph and its application to the beta model」（無向グラフに対するグレイバー基底とベータモデルへの応用）では、ランダムグラフの基本的な統計モデルであるベータモデルに対して、正確検定のための手法を構築している。その枠組みは Diaconis と Sturmfels (1998) のマルコフ連鎖モンテカルロ法に基づくものであるが、条件付き分布から観測され得るグラフに制約を伴うため、通常のマルコフ基底のみでは既約なマルコフ連鎖を構成できない。本章では、マルコフ基底を包含するグレイバー基底を用いることにより既約なマルコフ連鎖を与えることに成功している。グレイバー基底の構造はしばしば非常に複雑であるが、ここでは無向グラフに付随するト

ーリックイデアルのグレイバー基底に対するグラフ理論的な特徴づけを与え、その特徴づけに基づいてグレイバー基底の元をランダムに生成するアルゴリズムを与えている。

第４章「Markov degree for fiber-configurations」（ファイバー配置族のマルコフ次元）では、所与の配置行列のそれぞれのファイバーを新たな配置行列ととらえ，それらからなる配置行列族に対するマルコフ基底の性質について議論している。まず、ファイバー配置族のマルコフ基底の次数の上限を、もとの配置のマルコフ複雑度によって与えている。そして、具体的ないくつかの配置行列に対して、ファイバー配置族のマルコフ基底の次数の実際の最大値や、もとの配置行列のマルコフ複雑度を求めて、両者を比較している。具体的な例の中には，輸送多面体に付随する配置行列のように，数学的に重要なものも含まれている。

第５章「Conditional maximum likelihood estimation of discrete statistical models」（離散的な統計モデルの条件つき最尤推定）ではトーリックモデルを基礎として、さらに興味あるパラメータに関するモデル構造を追加した形の統計モデルを考え、そのパラメータ推定の問題を扱っている。基礎とするトーリックモデルのパラメータの影響を、対応する十分統計量を所与とした条件付き尤度を考えることによって除去し、興味あるパラメータのみを条件付き最尤法により推定する手法について議論している。その際、条件付き尤度の規格化定数の計算が困難な問題となるが、ここでは規格化定数がA-超幾何多項式とよばれる多項式であることに注目し、A-超幾何多項式についてこれまで知られている結果を拡張することにより、様々な計算手法が適用できることを示している。特に二元分割表の独立モデルを基礎とする場合について、級数展開法、ホロノミック勾配法、差分版のホロノミック勾配法の３種類の規格化定数の計算式を与えている。

最後に第６章「Concluding remarks」（結語）では、本論文の成果を簡潔に纏めると共に、今後の研究課題を提示している。

以上を要するに、本論文は代数統計学の枠組みにおける条件付き検定と推定という二つの基本的なテーマについて、統計の応用上で重要ないくつかの具体的な問題について新たな結果を導き、数理情報学分野の発展に寄与した。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。