

学位論文題名

離散多変量解析の理論とその応用に関する研究

学位論文内容の要旨

統計的推測の理論の根幹をなす推定における最尤推定法、検定における尤度比検定は、確率分布間の擬距離である Kullback-Leibler のダイバージェンスに基づき構築されていると考えることができる。つまり、最尤推定法においては真のモデルと観測値の Kullback-Leibler のダイバージェンスを最小とする値を推定量（最尤推定量）とし、尤度比検定においては、真のモデル（帰無仮説）と観測値との差異を Kullback-Leibler のダイバージェンスを用いて測り、その大きさの帰無仮説の下での分布によって検定を行なっている。この Kullback-Leibler のダイバージェンスを用いその局所化によって母数に依存する確率分布の母数空間に Fisher の情報行列によりリーマン計量を導入することによって、統計的推測の理論へ幾何学にアプローチする方法が Rao によって提唱され近年めざましい発展をとげている。これに対して Kullback-Leibler 以外のダイバージェンスを統計的推測に導入する研究が行なわれている。代表的なものとして J-ダイバージェンス及びパワーダイバージェンスがあげられる。J-ダイバージェンスは Burbea と Rao によってエントロピー関数の Jensen-difference に基づく確率分布の擬距離として提案され、エントロピー関数に、 $\alpha = 1$ の場合として Shannon-エントロピー、 $\alpha = 2$ の場合として Gini-index を含む Havrda-Charvát による α -エントロピーを用いることによって次数 α の J-ダイバージェンスとして具体的に表現されるものである。さらに、Burbea と Rao は、この次数 α の J-ダイバージェンスの局所化により、確率分布の母数空間に α -エントロピー計量というリーマン計量を導入している。しかし、従来の研究においては J-ダイバージェンスに基づく具体的な統計的推測は、Rao、Nayak 等の研究があるにすぎない。さらに尤度比検定、最尤推定に代表される Kullback-Leibler のダイバージェンスに基づく統計的推測方法と J-ダイバージェンスに基づく統計的方法との比較研究はほとんど行なわれていない。一方パワーダイバージェンスは Cressie と Read によって提案された Kullback-Leibler ダイバージェンスを $a = 1$ の場合として含む径数 a に依存するダイバージェンスの族である。このダイバージェンスの族から構築される検定統計量は J-ダイバージェンスと異なり大標本においてはすべての a について同等な統計量となり、漸近的検出力はすべて等しいと評価される。さらに、小標本での適合度検定における検出力の比較も彼らによって研究されている。しかし、応用において最も重要な分割表における一様性及び独立性の検定における小標本での検出力に関する研究は行なわれていない。

本論文はこれら J-ダイバージェンス及びパワーダイバージェンスに基づく多項母集団における統計的推測を明確化しその評価を与えることを目的とする。本論文は以下の7章より構成される。

第1章では、本研究の背景及び本論文の構成を示す。

第2章では、次数 α の J-ダイバージェンスに基づく多項分布の母数空間の幾何学的構造について論じ、この結果次数 α の J-ダイバージェンスの局所化によってリーマン計量を導入した n 項分布の母数空間がある n 次元ユークリッド空間の超曲面として具体的に表現されることが示された。このことにより J-ダイバージェンスを用いて統計的推測を行なう上での距離構造が明確となった。

第3章では、2章で与えられた構造を踏まえて、多項母集団の適合度における単純仮説及び複合仮説検定という具体的な問題に対して、次数 α の J-ダイバージェンスに基づく検定統計量を与え帰無仮説及び局所対立仮説のもとでの漸近分布を導出した。さらに、これらの統計量と対数尤度比統計量との漸近的検出力の比較を行ない対立仮説の方向によって対数尤度比統計量より検出力の高い検定統計量を与える α の値が存在することを示した。一方推定問題において、次数 α の J-ダイバージェンスに基づく未知パラメータの推定量を考察しその推定量の漸近分布を導出した。また、第2章で論じられた幾何学的構造によってこの推定量の解釈を行なった。

第4章では、複数の多項母集団の一様性検定に対して、次数 α の一般化された J-ダイバージェンスに基づく統計量を与え、帰無仮説及び局所対立仮説のもとでの漸近分布を導出した。この統計量に関して対数尤度比統計量との漸近的検出力の比較を行ない、対立仮説の方向によって尤度比統計量より検出力の高い検定統計量を与える α の値が存在することを示した。さらに、この検定方式の多元配置問題への拡張を行なった。

第5章では、複数の多項母集団の一様性検定に対して、パワーダイバージェンスに基づく統計量の局所的でない（固定された）対立仮説のもとでの分布の近似を統計量の漸近正規性に基づく方法及び統計量の展開式における2次までの項に基づく方法で行なうことによって、統計量の検出力近似を与えた。これらの検出力近似と、確率モデルからの直接計算によって得られた検出力との比較を行なった結果、展開式における2次までの項に基づく近似は小標本においても良い近似となっていることが示された。さらに、この近似を用いてパワーダイバージェンスに基づく2つの母集団の一様性検定統計量の検出力比較を行ない、対立仮説に対する各統計量の選択基準を考察した。

第6章では、分割表の独立性検定に対して、パワーダイバージェンスに基づく統計量の局所的でない（固定された）対立仮説のもとでの分布の近似を統計量の漸近正規性に基づく方法及び統計量の展開式における2次までの項に基づく方法で行なうことによって、統計量の検出力近似を与えた。これらの検出力近似と、モンテカルロシミュレーションによる検出力との比較を行なった結果、展開式の2次までの項に基づく近似は小標本においても良い近似となっていることが示された。さらに、パワーダイバージェンスに基づく統計量の検出力比較を行ない統計量選択の指針を示した。

第7章では、本研究で得られた結果をまとめ、今後の課題を検討した。

学位論文審査の要旨

主 査 教 授 佐 藤 義 治
副 査 教 授 新 保 勝
副 査 教 授 伊 達 惇

学位論文題名

離散多変量解析の理論とその応用に関する研究

統計的推測理論の推定における最尤推定法、検定における尤度比検定は、確率分布間の擬距離である Kullback-Leibler のダイバージェンスに基づき構築されていると考えることができる。さらに近年、Kullback-Leibler 以外のダイバージェンスを統計的推測に導入する研究が行なわれている。代表的なものとして C. R. Rao と J. Burbea により提案された J-ダイバージェンス及び N. Cressie と T. R. C. Read により提案されたパワーダイバージェンスなどが挙げられる。しかし、従来の研究においては、応用上重要な統計モデルに対する統計的推測、さらには検定における検出力に代表される推測方法の良さの評価はほとんど行なわれていない状況である。

本論文はこれら J-ダイバージェンス及びパワーダイバージェンスに基づく多項母集団における統計的推測方式を展開し、さらにその評価を行なっている。

第一章では、本研究の背景および本論文の構成を示している。

第二章では、次数 α の J-ダイバージェンスに基づく多項分布の母数空間における幾何学的構造について論じ、この結果次数 α の J-ダイバージェンスの局所化によってリーマン計量を導入した n 項分布の母数空間が、 n 次元ユークリッド空間のある超曲面として具体的に表現されることを示している。このことにより J-ダイバージェンスを用いて統計的推測を行なううえでの距離構造を明らかにしている。

第三章では、多項母集団の適合度における単純仮説及び複合仮説検定に関して、次数 α の J-ダイバージェンスに基づく検定統計量を与え、帰無仮説及び局所対立仮説のもとでの漸近分布を導出している。さらに、これらの統計量と対数尤度比統計量との漸近的検出力の比較を行ない対立仮説の方向によって対数尤度比統計量より検出力の高い検定統計量を与える α の値が存在することを示している。一方、推定問題において、次数 α の J-ダイバージェンスに基づく未知パラメータの推定量を考察しその推定量の漸近分布を導出している。

第四章では、複数の多項母集団の一様性検定に関して、次数 α の一般化された J-ダイバージェンスを用いたと解釈されるディスパージョン分析に基づく統計量を与え、帰無仮説及び局所対立仮説のもとでの漸近分布を導出している。この統計量に関して対数尤度比統計量との漸近的検出力の比較を行ない、対立仮説の方向によって尤度比統計量より検出力の高い検定統計量を与える α の値が存在することを示している。さらに、この検定方式の多元配置問題への拡張を行なっている。

第五章では、複数の多項母集団の一様性検定に関して、パワーダイバージェンスに基づく統計量の局所的でない、ある固定された対立仮説のもとでの分布の近似を統計量の漸近正規性に基づく方法及び統計量の展開式における 2 次までの項に基づく方法で行なうことによって、統計量の検出力近似を与えている。これらの検出力近似と、確率モデルから直接計算によって得られた検出力との比較を行なった結果、展開式における 2 次までの項に基づく近似は小標本においても良い近似となっていることが示されている。さらに、この近似を用いてパワーダイバージェンスに基づく 2 つの母集団の一様性検定統計量の検出力比較を行ない、対立仮説に対する各統計量の選択基準を考察している。

第六章では、分割表の独立性検定に関して、パワーダイバージェンスに基づく統計量のある固定された対立仮説のもとでの分布の近似を統計量の漸近正規性に基づく方法及び統計量の展開式における 2 次までの項に基づく方法で行なうことによって、統計量の検出力近似を与えている。これらの検出力近似と、モンテカルロシミュレーションによる検出力との比較を行なった結果、展開式の 2 次までの項に基づく近似は小標本においても良い近似となっていることが示されている。さらに、パワーダイバージェンスに基づく統計量の検出力比較を行ない統計量選択の指針を示している。

第七章では、本研究で得られた結果をまとめ、今後の課題の検討を行なっている。

これを要するに、著者は、離散多変量分布において代表的な多項分布に従う母集団に関する新たな推測方式及びその評価方式を展開し、統計情報処理において有益な新知見を得たものであり、データ解析学、情報処理学の発展に貢献するところ大なるものがある。

よって著者は、北海道大学博士（工学）の学位を授与される資格あるものと認める。