

氏 名 Xiaolei LU

学位(専攻分野) 博士(統計科学)

学位記番号 総研大甲第 1922 号

学位授与の日付 平成29年3月24日

学位授与の要件 複合科学研究科 統計科学専攻
学位規則第6条第1項該当

学位論文題目 Simultaneous confidence bands and the volume-of-tube
method

論文審査委員 主 査 教授 藤澤 洋徳
教授 栗木 哲
教授 福水 健次
教授 高 魏 Northeast Normal University

論文の要旨

Summary (Abstract) of doctoral thesis contents

This study focuses on simultaneous confidence bands and the volume-of-tube method. Simultaneous confidence bands have been used in various statistical problems. The volume-of-tube method can be used in the construction of simultaneous confidence bands.

The problem concerning the construction of simultaneous confidence bands in a regression model originates with Working and Hotelling (1929). They formalized this problem as the construction of confidence intervals for an estimated regression line, and provided a critical value by making use of the Cauchy-Schwarz inequality. Subsequently, many reports concerning the relaxation of these conditions have appeared.

In the case of one regression model, Wynn and Bloomfield (1971) pointed out that the use of the Cauchy-Schwarz inequality leads to conservative bands for simple regression and unrestricted domain of the explanatory variables. Uusipaikka (1983) constructed exact confidence bands for linear regression when the domain is a finite interval.

In the case of the linear regression models, there are a lot of research in literature. For example, Simultaneous confidence intervals are used in Scheffé (1953) to assess any contrasts between several normal means. In this study, the problem of assessing any contrasts between several simple linear regression models is considered by using simultaneous confidence bands. Using numerical integration, Spurrier (1999) constructed exact simultaneous confidence bands for all of the contrasts between several regression lines over the whole range $(-\infty, \infty)$ of the explanatory variable when the design matrices of the regression lines are all equal. Jamshidian, Liu, and Bretz (2010) proposed a simulation-based method to construct simultaneous confidence bands for all of the contrasts between the linear regression models when the explanatory variable is restricted to an interval and the design matrices of the regression lines may be different. Naiman (1986) gives a method for constructing conservative Scheffé-type simultaneous confidence bands for a single curvilinear regression model over finite intervals. Unlike these studies, we consider constructing simultaneous confidence bands for all of the contrasts between several nonlinear regression models. The tube formula is given in a mathematical form via the volume-of-tube method.

The chapters are arranged as follows. We provide a brief review of multiple regression models in Chapter 1. Chapter 2 summarizes simultaneous confidence bands for simple and multiple regression models, and we then

address the problem of the construction of simultaneous confidence bands for all of the contrasts between several nonlinear regression models. We propose simultaneous confidence bands of the hyperbolic type for the contrasts between several nonlinear (curvilinear) regression curves. Chapter 3 looks at the volume-of-tube method. We give the definition of the tube and critical radius, and then we summarize the volume-of-tube method for evaluating the upper tail probability. In addition, we discuss the expectation of the Euler-Poincaré characteristic heuristic. Moreover, we prove that the formula obtained is equivalent to the expectation of the Euler-Poincaré characteristic of the excursion set of the chi-square random process and, hence, is conservative. Using this result, Takemura and Kuriki (2002) provide an alternative proof that the confidence band of Naiman (1986) is conservative. Chapter 4 uses the volume-of-tube method to derive an upper tail probability formula for the maximum of a chi-square random process, which is sufficiently accurate in commonly used tail regions. The critical value of a confidence band is determined from the distribution of the maximum of a chi-square random process defined on the domain of the explanatory variables. The tube formula is given in a mathematical form. We prove that the simultaneous confidence bands we propose are conservative. This result is therefore a generalization of Naiman's inequality for Gaussian random processes. In order to test our method, we give a numerical example to determine the accuracy of the approximation formula we propose, which further demonstrate that the confidence bands obtained by the tube method are always conservative and very accurate. To investigate what happens under model misspecification, we conduct a Monte Carlo simulation study. The study shows, too small of a model should surely be avoided, whereas, a larger model has the disadvantage of having a wider confidence band.

As an illustrative example, the growth curves of consomic mice are analyzed in Chapter 5. A study under model misspecification is also conducted in the application. Chapter 6 considers the statistical parametric mapping approach as future work. Details of the proofs are in the Appendix.

Summary of the results of the doctoral thesis screening

提出された論文は、複数個の非線形回帰曲線や成長曲線のコントラスト（対比）に関する同時信頼区間の構成とそれを用いたデータ解析について論じたものである。英文で書かれており、全 6 章と Appendix, 計 54 頁からなる。

第 1 章は、本論文の序章である。研究の背景として、回帰分析に関する基礎事項の説明と、本論文全体の概略が述べられている。第 2 章は、同時信頼区間に関する基礎事項がサーベイされている。まず回帰分析におけるポイントワイズの信頼区間と同時信頼区間が、その違いとともに述べられている。また本論文で扱う複数個の回帰曲線のコントラストに関する同時信頼区間が、一般には解析的に構成することが難しいことが説明されている。ここでコントラストとは、和が 0 となるような係数に対する線形結合をいう。最後に、第 4 章で扱う問題が簡単に導入されている。第 3 章は、本論文で用いる数学手法的であるチューブ法 (tube method; volume-of-tube method) の要約に充てられている。チューブ法は、ガウス確率場の最大値の分布を精度良く近似するための積分幾何学的方法である。ここではその近似公式と、誤差評価の一般論が述べられている。またチューブ法の類似の手法であるオイラー標数法 (Euler characteristic heuristic) についても触れられている。続く第 4 章、第 5 章が本論文の主要部である。第 4 章では、第 2 章で導入した複数の非線形回帰曲線のコントラストに関する同時信頼区間を構成するために、ピボット (pivot, パラメータを含んだ統計量) を導入し、そのピボットのパラメータに関する最大値の分布を求めることによって同時信頼区間のバンド幅を決定できることを導いている。ピボットは、ガウス確率場とみなすことができるため、第 3 章で説明したチューブ法が適用可能となる。チューブ法の一般論に沿って近似公式を導き、数値計算例によってそれが正しいものであることを確認している。さらに、提案法が次元の添字集合をもつカイ 2 乗確率場に対するオイラー標数法近似と同等であることに着目し、提案近似公式が保守的であることを導いている。これはいわゆる Naiman (1986, Ann. Statist.) の不等式の一般化といえることができる。またモデル誤特定の下において、チューブ法により構成された同時信頼区間がどのような挙動をするかを、シミュレーションにより調べている。第 5 章は、第 4 章の理論的結果を用いたデータ解析例である。ここで用いているデータは国立遺伝学研究所のコンソミックマウスの生後 2~20 週の経時体重データである。標準マウス C57BL/6 (B6) を含む複数群のコンソミックマウスについて、時間と群間のコントラストの両面について同時信頼区間を構成し、多重性調整の下でも時点とコントラストの組み合わせによっては、有意差が検出されることを示している。第 6 章は、本研究に関連した今後の研究課題、Appendix は主として第 4 章で与えられた結果の証明の詳細である。

扱いが難しいとされてきた複数の非線形回帰曲線のコントラストに関する同時信頼区間構成を、シミュレーションではなく解析的に与えていること、またさらに与えた近似の保守性を示したことは評価できる。また実データ解析によって、提案手法の実用例を例示している点も評価できる。以上の状況に基づき、審査委員会は、本論文が統計科学の博士論文として十分な内容を有していると判断した。第 4 章と第 5 章の内容は査読付国際学術

(別紙様式 3)

(Separate Form 3)

雑誌 *Journal of Multivariate Analysis* に掲載されている。

総合研究大学院大学複合科学研究科における課程博士及び修士の学位の学位授与に係る論文審査等の手続き等に関する規程第 10 条に基づいて、口述による試験を実施した。口述による試験を実施した結果、出願者はその博士論文を中心としてそれに関連がある専門分野、その基礎となる分野及び英語の能力について博士（統計科学）の学位の授与に十分な学識を有するものと判断し、合格と判定した。