

令和 3 年 6 月 16 日現在

機関番号：14201

研究種目：基盤研究(C) (一般)

研究期間：2017～2020

課題番号：17K00054

研究課題名(和文) 不完全情報を伴う事象時間データのセミパラメトリック推測とその多次元化

研究課題名(英文) Semiparametric inferences for time-to-event data with incomplete data and their multidimensional extensions

研究代表者

杉本 知之 (Sugimoto, Tomoyuki)

滋賀大学・データサイエンス学部・教授

研究者番号：70324829

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：相関構造をもつ2変量のログランク統計量の漸近分布を定式化して、コピュラ型相関をもつ2変量事象時間データへの推測に応用した。セミ競争リスク問題の推測の研究、群逐次型の2変量ログランク統計量による標本サイズ的设计に応用する研究について一定の成果を得た。相対生存関数のSurvival回帰木によるセミパラメトリック推定法の研究を行って、回帰木の予測性能を測るための指標として、相対生存率モデル用のBrierスコアも提案して学会発表を行った。イベント時間とカレンダー時間を分離して表現する2変量ハザード・モデルのセミパラメトリック推測の研究も行い、そのセミパラメトリック推測のための計算手法を検討した。

研究成果の学術的意義や社会的意義

医療の世界では、がんや循環器疾患などで治療効果を測るために、事象時間データの分析は必須である。経済分野では倒産などのイベントを分析したり、製造業分野では在庫がなくなるまでの時間を分析したり、事象時間データの分析の適用例は多くみられる。そのようなデータの分析方法について、Cox回帰モデルなどの生存時間データの統計解析法は必須であり、その当該分野において、現在まで得られている統計的な分析方法、理論、計算手法を発展させるための研究を行い、一定の成果を得ることができたことは、学術的意義をもつ。さらに、これらの手法を実際のデータに応用していくことで、社会的に還元をなすことができる。

研究成果の概要(英文)：We formulated the asymptotic distribution of correlated bivariate log-rank statistics and applied it to inference for bivariate event-time data with copula-type correlation. Some results were obtained in the studies of inference for the semi-competitive risk problem, and application of group-sequential bivariate log-rank statistics to sample size design. We also studied a semi-parametric estimation method using survival regression trees for relative survival, and proposed a Brier score for relative survival models to measure the predictive performance of regression trees, which was presented at a conference. We studied semi-parametric estimation of bivariate hazard models in which event-time and calendar-time are represented separately, and examined computational methods for the semi-parametric estimation.

研究分野：統計科学

キーワード：生存分析 事象時間データ 回帰分析 多変量分布

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

Cox 回帰の推測の基礎理論の整備とともに、実地の問題に適合するように、データの守備範囲の拡大や、モデル自身の拡張がなされてきている。これらの発展を鑑みると、実地における有意義な展開を与える Cox 型計数過程モデルを定式化しても、例えば、潜在構造のモデル化に伴い、その想定モデルのあてはめには、不完全データを伴う形で推測しなければならない問題を伴うことが多い。不完全データを伴う Cox 回帰のあてはめの典型的な例を考えると、区間打ち切りデータ、傾向スコアを用いる場合も含め欠測共変量などがある。Cox 型計数過程モデルの推測において、不完全観測のもとで生じる諸種の部分尤度法には、それぞれに、独自の役割と用途をもつことが多い。本研究期間中に、これらの結果を、多変量生存時間モデルに発展させていくが、現状では、多変量生存時間モデルでは相関構造の定式化に、多変量正規分布のような便利な構造が利用しづらいため、その領域の統計理論や方法の研究はまだ解決できていない問題が多いというのが当初の背景にあった。

2. 研究の目的

本研究目的は、これまで Cox 型計数過程モデルの推測に関して得られた知見を、多変量生存時間モデルに発展させて、セミパラメトリック・周辺プロフィール尤度、周辺部分尤度、疑似部分尤度といった方法のそれぞれの役割と用途についての知見をより確固たるものにしていくことである。方法論の開発では、まず、コピュラ型の相関構造をもつ 2 変量事象時間データに対して、セミ競合リスク問題の推測の研究を精緻化して発展させる。この推定論では、EM アルゴリズムを用いて、死因既知のノン競合リスク問題に帰着し SPL を構成でき、群逐次 2 変量ログランク統計量による検定問題、さらに情報寄与ありの打ち切りに対する多変量ログランク統計量の修正版を研究する。相対生存モデルの研究では、相対生存関数の Survival 回帰木によるセミパラメトリック推定法の研究を行う。とくに、ここでは作られた回帰木の予測性能を測ることが必要であり、既存の指標に加えて、相対生存率モデル用の Brier スコアも提案する。これらの方法論では、標準誤差の計算のために、高次元でも計算効率の良い逆行列公式を利用し、周辺プロフィール尤度による推測を基本とする。ただし、周辺プロフィール尤度の解法の反復計算に問題が生じるときは、周辺部分尤度からの解法も求めて、両者のハイブリッド化を図る。また、社会変化をとり入れ経済的な分析にも拡張できるように、イベント時間とカレンダー時間を分離して表現する 2 変量ハザード・モデルのセミパラメトリック推測の研究も行う。重要かつ有意義な統計モデルとして、治癒混合モデル、区間中途打ち切りデータの推測、欠測共変量を伴う Cox 回帰モデルなどをとり上げて、3 タイプ尤度に基づく有用な理論と方法論の開発を行うことを研究の目的とする。

3. 研究の方法

本研究に関連する文献のレビューを徹底し、研究内容の計算プログラム開発を実装させる。多変量生存時間モデルをとりあげて、統計的推測の理論と応用の研究を遂行する。コピュラ型の 2 変量モデルからスタートして、セミ競合リスク問題や、相対生存モデルの方法論の開発を行い、セミパラメトリック推測のための 3 タイプの部分尤度の相互関係を活かした方法論の開発を行う。また、ハザード関数上の超高次元 Laplace 近似や、マルチンゲール分解の数理的方法を研究し、各尤度からの推定量の性質を、漸近論および有限標本挙動の両側面から調査する。いくつかの Cox 型計数過程モデルに対して、因果推論観察研究のためのセミパラメトリック推測への応用を視野に入れ、安定した最適化計算法の開発、有用な 2 標本検定法の研究、計算可能な漸近分散公式と区間推定法の導出などの研究を行いながら、多次元化の基盤整備を行う。

4. 研究成果

2 変量ログランク統計量の数理統計学的構造を調べ、2 変量生存時間関数の性質との接合を考えることで、通常マルチンゲール中心極限定理を異なる二つの時間軸に拡張することに成功し、相関構造をもつ 2 変量ログランク統計量の漸近分布を定式化することができた。コピュラ型の相関構造をもつ 2 変量事象時間データへの推測に応用した。まずは、セミ競合リスク問題の推測の研究を精緻化して発展させたこと、群逐次型の 2 変量ログランク統計量による検定問題、標本サイズの設計に応用する研究など行って、一定の成果を得ることができた。さらに情報寄与ありの打ち切りも可能にするような多変量ログランク統計量の修正版を研究して学会発表などを行った。相対生存モデルの研究では、相対生存関数の Survival 回帰木によるセミパラメトリック推定法の研究を行って、回帰木の予測性能を測るための指標として、相対生存率モデル用の Brier スコアも提案して学会発表を行った。社会変化をとり入れ経済的な分析にも拡張できるように、イベント時間とカレンダー時間を分離して表現する 2 変量ハザード・モデルのセミパラメトリック推測の研究も行い、そのセミパラメトリック推測のための、計算手法を検討した。Cox 回帰解析と層別解析は、両者の分析の目的は違うが、局外パラメータを排除して、関心パラメータについての推測を行う構造には共通点が多く密接に関係している。その知見を利用して、

層別解析の分析方法の研究を行い，計算代数統計を利用したマルコフ連鎖モンテカルロ法によるログランク検定の正確計算，ランダム効果モデルに対する DerSimonian-Laird 法の正確分布の研究を行い，学会発表を行うなどの一定の成果を得た．

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 3件/うちオープンアクセス 0件）

1. 著者名 Sugimoto T, Hamasaki T, Evans SR, Halabi S	4. 巻 26
2. 論文標題 Group-sequential logrank methods for trial designs using bivariate non-competing event-time outcomes	5. 発行年 2020年
3. 雑誌名 Lifetime Data Anal.	6. 最初と最後の頁 266 - 291
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s10985-019-09470-4	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Yokoyama S, Hamada T, Higashi M, Matsuo K, Maemura K, Kurahara H, Horinouchi M, Hiraki T, Sugimoto T, Akahane T, Yonezawa S, Kornmann M, Batra SK, Hollingsworth MA, Tanimoto A.	4. 巻 26
2. 論文標題 Predicted Prognosis of Pancreatic Cancer Patients by Machine Learning	5. 発行年 2020年
3. 雑誌名 Clinical Cancer Research	6. 最初と最後の頁 2411 - 2421
掲載論文のDOI（デジタルオブジェクト識別子） 10.1158/1078-0432.CCR-19-1247	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Sugimoto Tomoyuki, Hamasaki Toshimitsu, Evans Scott R., Sozu Takashi	4. 巻 36
2. 論文標題 Sizing clinical trials when comparing bivariate time-to-event outcomes	5. 発行年 2017年
3. 雑誌名 Statistics in Medicine	6. 最初と最後の頁 1363 ~ 1382
掲載論文のDOI（デジタルオブジェクト識別子） 10.1002/sim.7225	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計10件（うち招待講演 0件/うち国際学会 3件）

1. 発表者名 花田圭佑・杉本知之
2. 発表標題 試験数が少ない場合のランダム効果メタアナリシスの非漸近的性質
3. 学会等名 日本計算機統計学会第33回シンポジウム
4. 発表年 2019年

1. 発表者名 Hanada K, Sugimoto T
2. 発表標題 Non-Asymptotic Properties and Behaviors for Random-Effects Meta-Analyses When the Number of Studies Is Small
3. 学会等名 The 6th International Symposium on Biopharmaceutical Statistics (国際学会)
4. 発表年 2019年

1. 発表者名 Mizuma K, Sugimoto T
2. 発表標題 Log-rank test and its handicap procedure using computational algebraic statistics
3. 学会等名 The 6th International Symposium on Biopharmaceutical Statistics (国際学会)
4. 発表年 2019年

1. 発表者名 花田圭佑, 杉本知之
2. 発表標題 ランダム効果層別DL分析法の正確分布
3. 学会等名 日本計算機統計学会第32回大会
4. 発表年 2018年

1. 発表者名 永久保太士, 杉本知之, 後藤昌司
2. 発表標題 繰り返し測定値の順位に基づく解析: 2 群2 時点の場合
3. 学会等名 日本計算機統計学会第32回大会
4. 発表年 2018年

1. 発表者名 花田圭佑, 杉本知之
2. 発表標題 ランダム効果 DL 層別分析における標本分布
3. 学会等名 日本計算機統計学会第32回シンポジウム
4. 発表年 2018年

1. 発表者名 杉本知之
2. 発表標題 メタ分析P値とその挙動について
3. 学会等名 2017年度統計関連学会連合大会
4. 発表年 2017年

1. 発表者名 杉本知之・福井敬祐・伊藤ゆり
2. 発表標題 過剰ハザードモデルにおけるBrierスコアの推測
3. 学会等名 日本計算機統計学会第31回シンポジウム
4. 発表年 2017年

1. 発表者名 上田一輝・杉本知之
2. 発表標題 セミ競合リスクのもとでのコピュラ型2変量生存関数の推測
3. 学会等名 日本計算機統計学会第31回シンポジウム
4. 発表年 2017年

1. 発表者名 Hamasaki T, Evans SR, Sugimoto T, Asakura K
2. 発表標題 Designing multi-stage clinical trials with multiple event-time endpoints
3. 学会等名 Joint Statistical Meetings 2017 (国際学会)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

https://www.sci.kagoshima-u.ac.jp/sugimoto/appli1.html メタ分析 p 値とその挙動について http://www.jfssa.jp/taikai/2017/table/program_detail/pdf/201-250/10216.pdf 過剰ハザードモデルにおけるBrierスコアの推測 https://ci.nii.ac.jp/naid/40021426203
--

6. 研究組織		
氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------