

Title	Survival CART and its Applications
Author(s)	衛藤, 俊寿
Citation	大阪大学, 2004, 博士論文
Version Type	
URL	<a href="https://hdl.handle.net/11094/45034">https://hdl.handle.net/11094/45034</a>
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉</a> 大阪大学の博士論文について <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈/a〉</a> をご参照ください。

***Osaka University Knowledge Archive : OUKA***

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名	え とう とし ひさ 衛 藤 俊 寿
博士の専攻分野の名称	博 士 (工 学)
学位記番号	第 18822 号
学位授与年月日	平成 16 年 3 月 25 日
学位授与の要件	学位規則第 4 条第 1 項該当 基礎工学研究科情報数理系専攻
学位論文名	Survival CART and its Applications (生存時間研究における樹木構造接近法とその応用)
論文審査委員	(主査) 教授 後藤 昌司 (副査) 教授 白旗 慎吾 教授 稲垣 宣生

#### 論 文 内 容 の 要 旨

癌研究に代表される生存時間研究の実地の場合では、死亡のリスクに関与する予後因子を探索・評価することが、患者に効果的な治療を効率よく施すという見地から非常に重要である。従来、このような予後因子の探索・評価には、比例ハザード・モデルが汎用されてきたが、樹木構造接近法は、モデルに課される仮定や制約を殆ど必要とせず、しかも結果が視覚的に表現できることから、近年、生存時間解析の場で積極的に用いられる傾向にある。

生存時間解析の場における樹木構造接近法として、従来から 2 分岐法が適用されているが、多分岐法は提案されていない。実際に、生存時間解析の場で 2 分岐法のみによって予後因子を探索し、その結果を解釈することには予後因子の多様な可能性を表現するのに難がある。また、2 分岐法の結果の妥当性を診断するのに多分岐法が必要である。本稿では、このような状況に鑑み、生存時間解析の場面での多分岐法を開発する。

本論文では、最初に生存時間解析における中途打ち切りデータと生存率推定の方法を述べ、予後因子探索のための比例ハザード・モデルを紹介する。比例ハザード性を点検する方法として、対数累積ハザードプロットを紹介し、比例ハザード性を満たしていない共変量に比例ハザード・モデルをあてはめることの危惧と対処を論じる。また、ここで用いるデータについてその収集された経緯と目標を与える。

次に、樹木構造接近法概念を説明し、その開発の経緯を概括する。また、樹木構造接近法の特徴と適用範囲を明確にするために、樹木構造接近法を 6 種の特徴で類型化し、代表的な方法について各類型に対応づける。とくに、CART、AID、CHAID、CRUISE の諸法について、その概要とアルゴリズムを紹介し、数値例でその特徴を具体的に明示する。

予後因子の探索における樹木構造接近法について、2 分岐法の概要を論じ、RECPAM 法と 2 標本順位検定統計量を利用する方法を提示する。これらの方法について、いくつかの事例を紹介し、2 分岐法における問題点と多分岐法の位置づけに言及する。

多分岐法における分岐評価基準として採用した一般化 Kruskal-Wallis 検定統計量の概要と多分岐アルゴリズムを提示し、多分岐法におけるいくつかの留意点を述べる。2 分岐法で評価した事例を多分岐法で再解析し、その結果を 2 分岐法の結果と比較する。さらに、シミュレーションによって多分岐法と 2 分岐法の特徴を評価する。このとき、生存時間研究の特殊性、すなわち、中途打ち切り観測値を考慮してシミュレーションを実施する。その結果、多分岐

法は2分岐法の代替およびその妥当性評価に有用であることを示す。

最後に、本論文で研究・開発した多分岐法の残された問題点について触れ、今後の研究方向について述べる。

## 論文審査の結果の要旨

癌研究に代表される生存時間研究の実地の場では、死亡のリスクに関与する予後因子を探索・評価することが、患者に効果的な治療を効率よく施すという見地から非常に重要である。従来、このような予後因子の探索・評価には、比例ハザード・モデルが汎用されてきたが、樹木構造接近法は、モデルに課される仮定や制約を殆ど必要とせず、しかも結果が視覚的に表現できることから、近年、生存時間解析の場で積極的に用いられる傾向にある。

生存時間解析の場における樹木構造接近法として、従来から2分岐法が適用されているが、多分岐法は提案されていない。実際に、生存時間解析の場で2分岐法のみによって予後因子を探索し、その結果を解釈することには予後因子の多様な可能性を表現するのに難がある。また、2分岐法の結果の妥当性を診断するのに多分岐法が必要である。本稿では、このような状況に鑑み、生存時間解析の場面での多分岐法を開発する。

本論文では、最初に生存時間解析における中途打ち切りデータと生存率推定の方法を述べ、予後因子探索のための比例ハザード・モデルを紹介する。比例ハザード性を点検する方法として、対数累積ハザードプロットを紹介し、比例ハザード性を満たしていない共変量に比例ハザード・モデルをあてはめることの危惧と対処を論じる。また、ここで用いるデータについてその収集された経緯と目標を与える。

次に、樹木構造接近法を説明し、その開発の経緯を概括する。また、樹木構造接近法の特徴と適用範囲を明確にするために、樹木構造接近法を6種の特徴で類型化し、代表的な方法について各類型に対応づける。とくに、CART、AID、CHAID、CRUISEの諸法について、その概要とアルゴリズムを紹介し、数値例でその特徴を具体的に明示する。

予後因子の探索における樹木構造接近法について、2分岐法の概要を論じ、RECPAM法と2標本順位検定統計量を利用する方法を提示する。これらの方法について、いくつかの事例を紹介し、2分岐法における問題点と多分岐法の位置づけに言及する。

多分岐法における分岐評価基準として採用した一般化Kruskal-Wallis検定統計量の概要と多分岐アルゴリズムを提示し、多分岐法におけるいくつかの留意点を述べる。2分岐法で評価した事例を多分岐法で再解析し、その結果を2分岐法の結果と比較する。さらに、シミュレーションによって多分岐法と2分岐法の特徴を評価する。このとき、生存時間研究の特殊性、すなわち、中途打ち切り観測値を考慮してシミュレーションを実施する。その結果、多分岐法は2分岐法の代替およびその妥当性評価に有用であることを示す。

最後に、本論文で研究・開発した多分岐法の残された問題点について触れ、今後の研究方向について述べる。

ここで提示された諸種の統計的方法論はいずれも、実地を意識した形式で展開が諮られている点で、相当に魅力的であり、広範な適用可能性を秘めている。よって本論文を博士(工学)の学位論文として価値のあるものと認める。