

氏名 貝 瀬 徹

学位（専攻分野） 博士（学術）

学位記番号 総研大甲第166号

学位授与の日付 平成8年3月21日

学位授与の要件 数物科学研究科 統計科学専攻

学位規則第4条第1項該当

学位論文題目 Survival Analysis Based on The Information
Criterion EIC

論文審査委員 主査教授 北川源四郎

教授 石黒真木夫

助教授 佐藤俊哉

教授 鎌倉稔成（中央大学）

副部長 松原敏樹

（癌研究会附属病院）

論文内容の要旨

生存時間解析は、医学、工学、経済学などの広い分野において寿命データから生存率を推定する目的で使われている。この場合、想定された生存率モデルに基づいて解析を検討するのが一般的である。このような生存率モデルとしてワイブル分布やガンマ分布などのパラメトリック・モデル、またカプラン-マイヤー (KM) モデルなどのノンパラメトリック・モデルが知られている。さらに、個体の持つ因子の影響を考慮する場合、パラメトリックの加速モデルやノンパラメトリックのコックス生存率モデルが用いられる。

さて、情報量規準 AIC は既に多くの統計解析において使われている。しかし、生存時間解析において AIC に基づく解析方法を提案した報告は少ない。AIC はパラメトリック・モデルに対して最尤法を前提に用いられるため、生存時間解析のようにノンパラメトリック・モデルを想定する場合、AIC の適用には困難さを伴う。ところが、近年 AIC を拡張した情報量規準 EIC が提案されている。EIC はブートストラップに基づいており、最尤法以外の推定法にも適用可能である。したがって、EIC によりノンパラメトリックでのモデル選択も可能であることが示唆される。また、生存率は個体の持つ状態や環境といった因子により大きく異なる。因子の影響は加速モデルやコックス生存率モデルにより表現することができるから、EIC によりこれらのモデルを評価することで因子と生存率の関係を解析することが予想できる。

以上のような背景から、本論文では EIC に基づく生存時間解析について提案した。ここではパラメトリックとノンパラメトリックのモデルを比較するため、線形化補間したカプラン-マイヤー (LIK) モデルと線形化補間したコックス生存率 (LIC) モデルを導入し、EIC の適用方法について明らかにした。このとき、ブートストラップ法として model-base の方法を用いた。これらの方法は、シミュレーションによりその有効性が示された。すなわち、提案する EIC に基づく解析により、パラメトリック、ノンパラメトリックの生存率モデルを統一的な立場で検討することが可能となった。特に、ノンパラメトリック・モデルは標本数に比例してバイアス (期待平均対数尤度と対数尤度の差) が大きくなるという特性を持つことが示された。このことにより、パラメトリックとノンパラメトリックのモデルの違いについて見通しのよい説明を与えることができた。さらに、EIC に基づいて生存率モデルを検討することで、生存率と因子との関係を明確にすることも可能となった。

提案する EIC に基づく解析方法の実証による有効性を示すため、胸部食道癌予後データを解析した。解析に用いたデータは 1985 年 6 月から 1992 年 10 月までの癌研病院外科による胸部食道癌予後データ、症例数 143, うち死亡データ 56, 生存打ち切りデータ 87 である。手術後の生存時間の他に、手術時点でのリンパ節転移個数、壁深達度、リンパ節転移パターンが各患者ごとに得られている。食道癌は手術時点の因子の状態により予後が大きく異なることが知られている。この解析では、これら予後因子と予後との関係を明らかにすること、また生存率モデルに基づいて 50 パーセント点を推定することを目的とした。まず、因子のカテゴリーによる層別全体はパラメータ数の多い制約された生存率モデルであることに着目し、EIC に基づいて単一因子の層別方法 (評価システム) の良さを検討した。ここでの生存率モデルは、LIK モデル、ワイブル分布、ガンマ分布を用いた。50 パーセント点の推定の場合、どの生存率モデルが適しているかを判断する必要があるが、ここでは

EICによるモデル比較を行なった。さらに複数の因子の影響を考慮する場合、層別だけによる解析では困難である。そこで、胸部食道癌予後データにおいて複数の因子の影響を考慮した解析を検討した。解析にはLIC生存率モデルとワイブル、ガンマ加速モデルを想定し、モデル選択にEICを用いた。EICを用いることで生存率に対する因子の影響の解析を行なった。解析結果として次のことが得られた。

1. 単一因子の評価システムとして、リンパ節転移個数が適していることがわかった。また、その次にリンパ節転移パターン、さらに癌の壁深達度の順であった。
2. この場合、ワイブル分布による50パーセント点の推定結果が適していた。
3. 複数の因子の影響を考慮した解析では、リンパ節転移個数、リンパ節転移パターン、癌の壁深達度の影響を回帰で表現したモデルが適していた。また、このモデルの方がリンパ節転移個数単一の評価システムより良いことが示された。
4. この場合、ワイブル加速モデルによる50パーセント点の推定結果が適していた。

以上の結果は医学的にも興味深いものであり、ここで提案するEICに基づく生存時間解析が実際の解析においても効果的であることが示された。

論文の審査結果の要旨

本委員会は貝瀬徹氏の論文について、数物科学研究科における課程博士の授与に係る論文審査等の手続き等に関する規程に基づき、公開の論文発表会を開催し審査を行なった結果、統計科学に関して学位を授与するに十分な内容と形式を備えているものと判断した。理由は以下の通り。

1) 論文の概要

本論文は情報量規準 EIC に基づく生存時間解析の手法を提案したものであり、全体の構成は6章からなる。

第1章では、生存時間解析についての背景と現状、そしてこの論文で取扱う問題の設定、特に打ちきりデータの取り扱いをめぐる問題が述べられている。

第2章では、この論文に関連する生存時間解析の基本的な概観が述べられ、パラメトリック生存率モデルを用いた解析法とノンパラメトリックな Kaplan-Meier 法の結果を情報量規準で評価・比較するため、線形化補間した Kaplan-Meier (LIK) モデルが導入されている。さらに共変量の影響を調べるための線形化補間した Cox 生存率 (LIC) モデルも導入されている。

第3章では、情報量規準 EIC の概要を述べ、第2章で導入した LIK モデルと LIC モデルに対しての EIC の適用手法が述べられている。とくに、モデルに基づくブートストラップ法の利用を提案し、その理由について説明している。

第4章では、生存率モデルにおける EIC の適用についてシミュレーションによりその有効性を示している。さらに、パラメトリックとノンパラメトリックの生存率推定法の違いについて情報量規準に基づいて説明を与えている。また、生存率曲線の共変量への依存の構造を EIC に基づいて解析する手法をシミュレーションによる例を用いて説明している。

第5章では、胸部食道癌の手術以後における生存時間のデータを EIC に基づいて解析している。解析結果として、手術時点での因子の生存率におよぼす影響と、生存率モデルとしてはパラメトリックモデルが適していることなどを明らかにしている。

第6章では本論文の結論が述べられている。

2) 論文の評価

本論文の貢献は、打ちきりを含むデータから生存率曲線を推定する問題に新しい情報量規準 EIC に依拠した方法を持ち込むことによって、パラメトリックモデルによる解析とノンパラメトリックな解析法を統一的な規準で比較することが可能であることを示した点にある。提案された方法を用いることによってパラメトリックモデルを用いることの妥当性の検証が可能となり、生存率曲線の共変量への依存のありかたを研究する際の客観的なモデル評価と選択が可能になった。

情報量規準 EIC に着目した本論文の提案は統計学において高く評価できるとともに医学、工学の分野においても貢献するところが大きいと判断される。