

学習アルゴリズムによる整合的多重脱退モデル推定

尾上辰徳* 松山直樹†

2017年1月31日投稿

2017年2月14日受理

概要

本研究は、保険数理で用いられる多重脱退モデル推定の基本問題を解消するための新たな手法（交互イタレーション）を提案するものである。交互イタレーションは公的統計による推移確率の不完全観測から出発して状態間制約と年齢間制約を充足する整合的多重状態モデルを推定する学習アルゴリズムであり、尤度の意味での収束性を示すことができる。また、粗い年齢群団数値しか得られない公的統計データを元データの再現性を確保しつつ各歳別に展開する手法（積分補間）も用いる。具体的に、がん疾病の多重脱退モデル推定に本手法を適用し、それらの有用性を示す。

キーワード：がん多重脱退、定常社会モデル、交互イタレーション、欠測値、公的統計

1. 多重脱退モデル推定問題

(1) 観測推移確率と多重脱退モデルの不整合

生命保険数理は、長期の保険期間に対応するために、状態推移確率が状態と年齢のみによって決まる定常社会モデルに基づき構成されているが、現実において定常社会の観測は不可能である。死亡のような一状態モデルでは現実の状態推移確率の観測値を定常社会モデルの状態推移確率とみなしたモデル構築が可能であるが、医療保険分野の多重脱退モデルにおいては根本的な不整合が存在する。すなわち、仮に状態推移確率の完全観測が可能だとしても、状態別の推移確率と全状態を統合した推移確率の関係は現実の非定常な状態別人口構成の影響を受けるため、定常社会モデルにおける状態別人口構成における関係と不整合となる。たとえば、 l_x, ℓ_x^j はそれぞれ定常社会モデルの年齢別人口、状態 j における年齢別人口を表し、 $\hat{l}_x, \hat{q}_x, \hat{\ell}_x^j, \hat{q}_x^j$ は現実の人口、死亡率、状態 j における人口、死亡率の観測値とすると、現実の非定常観測では、状態別の死亡数合計は全死亡数に一致するため $\hat{l}_x \hat{q}_x = \sum_j \hat{\ell}_x^j \hat{q}_x^j$ が成り立つが、定常社会モデルの人口構成では $l_x \hat{q}_x \neq \sum_j \ell_x^j \hat{q}_x^j$ となり不整合を生ずる。多重脱退モデルは、年齢間制約条件（年齢による漸化式が成り立つ）と状態間制約条件（ある年齢の状態別人口合計は全人口に一致する加法性が成り立つ）という二重の制約条件を満たすものでなければならないが、観測値を直接用いる限り、いずれかの制約条件は無視せざるを得ない。例えば、先行研究（友寄 [2015]）では年齢間制約条件は無視されている。

(2) 公的統計に依存する制約

医療保険分野の多重脱退モデル推定では状態数の増加に伴い状態別の標本数が少なくなるため、一般的には観測数が大きい国レベルの公的統計に依存せざるを得ない。公的統計の制約として、(i) 必

*明治大学大学院先端数理科学研究科

†明治大学総合数理学部 〒164-8525 中野区中野 4-21-1 ma2yama(at)meiji.ac.jp

本稿作成に当たり、実務的アドバイスを頂戴した住友生命の金村慶二氏と有益なコメントを頂戴した匿名査読者に感謝する。