

科学研究費助成事業 研究成果報告書

平成 30 年 6 月 4 日現在

機関番号：12301

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00435

研究課題名(和文)医療サービス履歴からの個人状態の時間発展の推定とその応用

研究課題名(英文) Estimation of temporal evolution of personal condition from medical service history and its application

研究代表者

関 庸一 (Seki, Yoichi)

群馬大学・大学院理工学府・教授

研究者番号：90196949

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：実証分析対象として、大阪府和泉市における平成23～平成28年度の医療サービス履歴データ(国民健康保険のレセプトデータ、および、特定健康審査結果)と、当該期間の運動指導プログラムの実施結果を利用した。特定健診データから自己組織化マップにより健康状態類型を抽出し、これとレセプトデータから、類型毎の主要病名、医療費を求めて疾病リスク等を明らかにした。さらに、6年間の健康状態変化を検討した結果から、メタボリックシンドロームに相当する個人類型は、医療費、受療率ともに他類型に比べ多いこと、運動指導プログラムは健康状態の改善に寄与することが示唆された。また、最尤法による自己組織化マップの拡張を検討した。

研究成果の概要(英文)：As an empirical analysis, we used the medical service history data (Receipt Data of National Health Insurance and Specific Medical Examination Result) of Izumi City, Osaka Prefecture, from FY2011 to FY2016, and data of the exercise instruction program in the period. From the specific medical examination data, the health condition type was extracted by the self-organizing map, and the major disease name and medical cost for each type were obtained using the receipt data, and the disease risk etc. of each type was clarified. In addition, as a result of considering the change in health condition type during the 6 years, it is suggested that medical expenses and rates of treatment acceptor of the types corresponding to metabolic syndrome is larger than the other types, and that the exercise instruction program contribute to improving health condition. We also studied development of self-organizing map by maximum likelihood method.

研究分野：統計科学

キーワード：データマイニング ビッグデータ解析 自己組織化マップ 状態遷移 最尤法 医療保険サービス 特定健康診査

1. 研究開始当初の背景

近年、情報収集技術の普及とともに、蓄積される情報は時空間上で稠密なものとなって来ている。つまり、多くの対象について長期間多面的、かつ、悉皆に収集できるようになってきている。このようなデータの中で、時系列上の特徴量の変化履歴のデータは、その中に個人についての多面的な知識を含むと考えられ、適切なモデルと分析方法が提供されれば、各種の意思決定支援の可能性は大きく広がる。しかし、これらの履歴データ全体を通じて、個体の履歴を総括的に追跡して分析した研究はない。たとえば、マーケティングや医療分野で既に多くの分析がなされているが、購買や死亡などの特定のエンドポイントへの影響評価など、データの特定の一面を抽出して分析するものに留まる。

本研究では、このようなデータから個人の長期的変化のモデル化、つまり、どのような個人はどのような状況においてどのような状態となるのかを明らかにする方法論を対象とする。

これら进行分析する際の困難さは、まず、履歴の多様さにある。データが他の業務などに用いるために収集されたトランザクションであり、分析のために計画して選別されたものではないため、背後に多様で異なった生成過程がある場合が多く、全データを一括したモデル化に無理があったからである。

申請者は、上述のようなデータに対し、平成 23~25 年度科学研究費：「時空間上の大規模履歴の自己組織化による状態遷移確率モデルの構築」および、それに至る一連の研究において、SOM (Self-Organizing Maps, 自己組織化マップ)(Kohonen, 2001) などを用い、個体集団の細分を作成するアプローチによる事例の蓄積を行って来ている。

2. 研究の目的

現在、医療・介護サービスの分野では、レセプトや健康診断・受診記録などの形で、時空間上での多次元の個人履歴データが、かなり稠密に蓄積されている。しかし、それらを突合し、一貫した履歴情報として理解するには、データ突合と対象の多様性把握の面に困難が存在している。

本研究では、この大量データに対し、整合的な履歴データ構築の方法論を与えるとともに、そこから個体状態のもつ位相的構造を可視化することで、多様な個人のサービス選択とサービス受給効果を統計的にモデル化する方法論を与える。これにより、公共的サービスなどの需要予測、サービスの推奨、適正なサービス提供など、多く場面で多様な個人のニーズに適合した意志決定を可能とすることを目的とする。

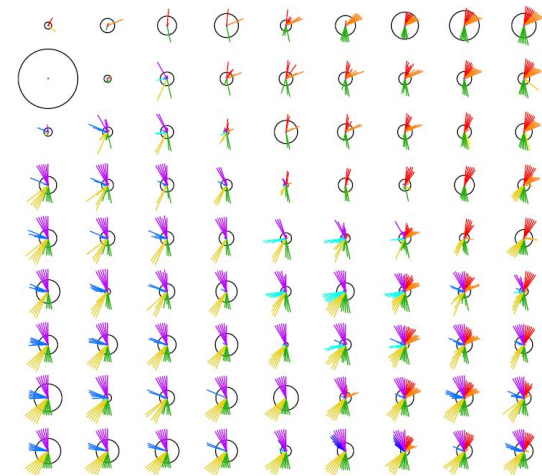
3. 研究の方法

(1) 代表的な個人履歴データとして、以下のデータ等を対象として研究を行った。

(a) 医療・介護保険の治療・サービス履歴 (レセプト/受給申請の健康保険/行政データなど)

(b) 健康診断・健康指導履歴(メタボ健診とスポーツ指導履歴データなど)

(2) 基礎とするデータ解析手法としては、SOM (Self-Organizing Maps, 自己組織化マップ)をもちいる。SOM は、平面上に格子状に配置されたノードごとに、代表類型(参照ベクトル)を求め、個体を最も類似した類型へ層別する方法である(図 1)。計算量がサンプル数 n に比例のオーダーであるため、多量サンプルデータにも適用可能である点が利点となる。



66 次元の利用の特徴量をもつクレジットカードの月々の利用状況を類型化した結果例。9×9 のノード配置に、参照ベクトルをレーダーチャートで示してある。

図 1: SOM のノードごとの参照ベクトル

(3) 履歴データに含まれる個人の状態を自己組織化することにより、自然な状態分割を実現し、医療サービス履歴データから個人の状態に関する状態遷移確率モデルを構築することを可能とする。

対象とする医療サービス履歴を分析する際の困難さは、まず、データの時点整合化である。多方面から収集されたデータは時点に関して不整合となる。これを時点整合的な特徴量化することが必要である。また、多様な個人の多様な状態を、推移を理解しやすいように層別する状態の分節化が必要である。さらに、個人の変化の多様性をどのようにモデル化するかに課題がある。

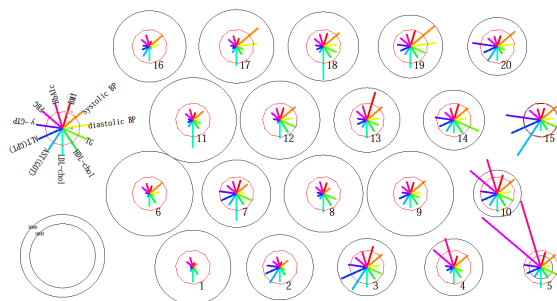
本研究では、以上に対し、これまでの研究成果を基礎として、時点整合化、状態の分節化、状態遷移確率の 3 点に注目して、広く適用可能な履歴モデルの構築法を検討した。

4. 研究成果

(1) 実証分析対象として、大阪府和泉市における平成 23~平成 28 年度の医療サービス履歴データ(国民健康保険のレセプトデータ、および、特定健康審査結果)と、当該期間の

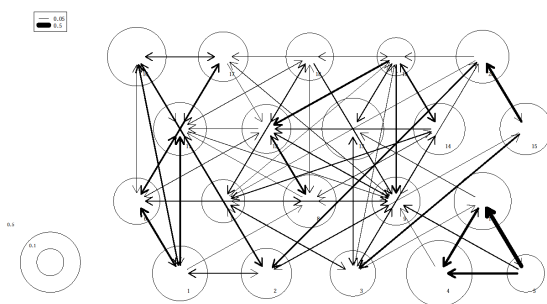
運動指導プログラムの実施結果を利用した。特定健診データから自己組織化マップにより健康状態類型を抽出し(図2)、これとレセプトデータから、類型毎の主要病名、医療費を求めて疾病リスク等を明らかにした。さらに、6年間の健康状態変化を検討した結果(図3)から、メタボリックシンドロームに相当する個人類型は、医療費、受療率ともに他類型に比べ多いこと、運動指導プログラムは健康状態の改善に寄与することが示唆された[学会発表]。

以上の分析にあたり、時点整合化のため、年度ごとの健診結果を年度の代表値として、年度の医療履歴との突合を行った。状態遷移の把握間隔としては、3年または1年間隔での変化に注目して、その特徴を検討した。時点をより詳細にみる突合方法と状態遷移確率のモデル化が今後の課題として残されている。



13 角形が基準上限値。円の面積がノード所属人数

図2 健康状態類型



5%以上の遷移のみ描画。円は自己遷移を表す。

図3 3年間隔での類型間遷移比率

(2) 個人状態の分節化にあたり、最尤法による自己組織化マップの拡張を検討した。従来の標準 SOM では、性格の異なったノードがマップ上で隣接するため、遷移関係などを的確に把握することが難しかった。そこで、研究代表者らは SOM の拡張法として、マップをデータの位相構造に適合させて変形する自己組織化として、FSOM (Flexible SOM) [1] などを検討してきた。しかし、FSOM では、マップ最適化のための統一的な基準が設定

できておらず、一貫したモデル選択が難しいことが課題として残されていた。そこで、SOM マップの各ノードにデータの確率分布モデルを用意し、サンプルのベクトルデータのノード確率分布での負の対数尤度を、サンプルとノードとの非類似度として、サンプルを最整合ノードに対応付け、各サンプルクラスターでノードの確率分布のパラメータを最尤推定することで確率モデルを構成する方法を提案した。このモデル化では、データの確率分布モデルを適切に設定することさえできれば、サンプルの最尤法による分節化が可能となり、さらに、MDL 基準などのモデル選択基準でクラスター数やマップ形状を評価選択できることとなる。マップの評価が可能となれば、マップの改善算法を検討できることとなる。この最尤モデルの計数値データの場合について、具体的な算法を提案した[学会発表]。さらに、一般分布の場合への拡張とノード間隣接関係の更新算法、および実データでの検証について、検討を進めた。

(3) 以上で得られた個人履歴の分析法等を活用し、経営科学系研究部会連合協議会主催のデータ解析コンペティションに参加して、消費者行動の大規模履歴 (ID 付 POS データ) について、SOM を用いた分析結果をまとめた[雑誌論文、学会発表]。

(4) 今後の展望としては、個人状態の分節化のための最尤自己組織化マップのモデルの拡張を、実データでの検証とともに進める予定である。

[1] サンプルの所属度に応じた可変自己組織化マップ, 多賀谷侑史, 安藤晋, 関庸一, 情報処理学会論文誌・数理モデル化と応用, Vol.5, No.3, pp.1-13(2012).

5. 主な発表論文等

〔雑誌論文〕(計1件)

川名純平, 諏訪竜也, 関庸一, 食品スーパーの多様性と顧客購買行動との関係分析, 日本オペレーションズ・リサーチ学会和文論文誌, 査読有, Vol.61, pp.23-44, 2018, http://www.orsj.or.jp/~archive/pdf/j_mag/Vol.61_J_023.pdf

〔学会発表〕(計5件)

Naoki Ikabata, Yoichi Seki, Jyunya Hujisaki, Takashi Sanada, Azusa Kamachi, Mika Tanigawa, Extraction of health condition type from personal histories of specific medical examination, 15th ANQ Congress, Sep 20-21, 2017, Kathmandu Nepal

関庸一, 川名純平, 諏訪竜也, 自己組織化マップによる大規模データ中の多様性の把握法, 日本経営工学会秋季大会予稿集, pp.146-147, 日本教育会館, 2016年10月

茂木亮祐, 関庸一, 最尤自己組織化マ

マップによる計数値ベクトルのセグメンテーション, 統計関連学会連合大会講演報告集, p.112, 金沢大学角間キャンパス, 2016年9月

五十畑直紀, 関 庸一, 藤崎 淳矢, 真田 崇, 蒲地 梓, 谷川 美香, 特定健康診査データからの健康状態タイプの抽出, 日本品質管理学会第110回研究発表会要旨集, pp.137-140, 日本科学技術連盟・東高円寺ビル, 2016年5月

茂木亮祐, 関 庸一, 最尤自己組織化マップによる来店・購買履歴からの顧客行動習慣の抽出, 日本経営工学会平成25年度春季研究大会予稿集, pp.18-19, 首都大学東京, 2015年5月

〔図書〕(計0件)

なし

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

なし

6. 研究組織

(1) 研究代表者

関 庸一 (SEKI, Yoichi)

群馬大学・大学院理工学府・教授

研究者番号: 90196949

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

なし