

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 7 日現在

機関番号：11501

研究種目：若手研究(B)

研究期間：2013～2014

課題番号：25730012

研究課題名(和文) 時空間点過程によるインフルエンザ伝播モデルの構築

研究課題名(英文) Spatio-temporal point process models for the spread of influenza in Japan

研究代表者

坂口 隆之 (Sakaguchi, Takayuki)

山形大学・教育文化学部・講師

研究者番号：10436496

交付決定額(研究期間全体)：(直接経費) 1,400,000円

研究成果の概要(和文)：インフルエンザの伝播を統計学的に説明するためのモデル構築に取り組んだ。対象地区は宮城県を選定した。宮城県は仙台市の他、7つの保健所管轄区域に分けられる。それぞれの区域ごとにインフルエンザ罹患数データが週報として時系列で公開されており、これらを利用した。モデルへの近隣情報の取り込みとして、メッシュデータのような保健所管轄区域とは異なる空間集積レベルのデータの利用について検討をおこなった。現状で推定したモデルはきわめて基礎的なものであるため、さらに人の往来などの情報をモデルに盛り込んで、より発展的な解析をおこなう必要がある。

研究成果の概要(英文)：The purpose of this study is to construct a model which explains the spread of influenza in Japan. The data collected in Miyagi prefecture were used. There are 8 healthcare centers in Miyagi prefecture. The weekly number of influenza patients reported from each healthcare center is available. We constructed a basic model for the data. It is helpful to use auxiliary variables, for example, information about the flow of people. However the use of auxiliary variables causes so-called Change of Support Problem. We investigated the possibility of using auxiliary variables observed at different spatial scales. In the future, the basic model needs to be developed by using auxiliary variables.

研究分野：統計学

キーワード：時空間点過程 インフルエンザ

1. 研究開始当初の背景

インフルエンザや結核、マラリア、あるいはエイズなどの感染症がどのように人々の間に蔓延し、そして終息していくのか、その機序を明らかにすることは疫学研究における重要なテーマである。実際、多方面からの研究がなされているが、その中には数理的な観点から伝播の仕組みを明らかにしようという試みもあり、主として微分方程式によるモデル化がおこなわれる(稲葉(2008))。

上記の微分方程式によるモデリングは、基本的に決定論的な方法であるといえる。それに対して、罹患者の時間的・空間的な拡がりや“ランダムな現象”と捉え、統計モデルで説明しようという動きもある。疫学現象の地理的な側面に着目し、これを空間データと捉えることで空間統計学的なアプローチをする分野は空間疫学と呼ばれている(丹後他(2007))。

空間統計学の主要理論の1つである空間点過程は、疫学分野においては疾病の集積性検出(高橋他(2008))や疾病患者の地理的分布に対する説明手段として利用されている(Best et al. (2000), Hohle(2009), Meyer(2009))。Best et al. (2000)は、喘息患者の空間分布を非定常 Poisson 点過程によってモデル化している。最近では、Hohle(2009)が感染症伝播の時間的な側面に注目し、これに対する時系列点過程モデルを提案している。また、Meyer(2009)は特定の感染症を想定しているというよりは、より一般的な枠組みで感染症伝播に対する時空間点過程モデルを与えている。

統計地震学で ETAS モデルが成功を収めているように、疾病の地理的側面を扱う空間疫学においても、空間点過程による積極的なモデリング研究が必要である。感染症伝播の特徴的な点は、人々の接触によって感染が起こり、さらに“感染者が移動”をすることで他地域にも感染が広がっていくということ

であろう。つまり、“人の流れ”が重要な役割を果たしていると考えられるため、人に関する情報をうまく反映させたモデル構築が必須である。

2. 研究の目的

本研究の目的は、統計数理モデル(時空間点過程モデル)を用いてインフルエンザの伝播を数学的に説明することである。

インフルエンザは人々の接触によって引き起こされるため、感染者がいつどこに存在するかという情報に加え、感染者がどの程度の人と接触するかという情報は重要だと考えられる。そして、人々の接触機会は人口や人の流れなどに相関すると思われる。人口や人の流れには地域差があるため、地理データとしてこれらの情報をモデルに取り込む必要がある。

以上を踏まえて、統計学的研究の観点からは、本申請課題は大きく分けて、以下の2つの目的を持っている。

(1) 空間点過程理論の応用研究の側面:

空間点過程モデルが感染症伝播をどの程度まで説明しうるか、その可能性を探る。

Hohle(2009)や Meyer(2009)は、空間点過程の条件付き強度という道具によって、感染症伝播のパターンを時空間的に記述している。この種の研究では、モデル化の対象となる現象の特徴をうまく条件付き強度関数に取り込むことが肝要である。当然ながら、感染症の種類が異なれば、それに応じて拡大メカニズムも変わると考えられる。

一般論および海外の実証的な研究は Paul et al. (2008), Meyer(2009), Paul and Held(2011)などによってなされている。そこで、本研究では日本におけるインフルエンザの伝播に焦点を当てて、時空間点過程による説明を試みる。

(2) Change of Support Problem(COSP)の側面:

あるデータが校区レベルで観測されていても、別のデータが市町村レベルで観測されていれば、これらのデータには観測スケールの違いがあるため、統合した解析の際には統計的なバイアスを生じうる。この問題は Change of Support Problem(COSP)と呼ばれる (Gotway and Young(2002))。本件においても、例えば、インフルエンザ罹患患者数データは保健所の管轄地区レベルで集計されているのに対し、国勢調査の人口データはメッシュレベルで集計されており、観測の空間スケールは異なる。このような観測スケールの違いを考慮し、どのようにモデルに取り込むべきかの検討をおこなう。

3. 研究の方法

対象地区は東北地方の中核である宮城県とする。宮城県には8つの保健所がある。それぞれの管轄地域ごとにインフルエンザ罹患患者数のデータが時系列(週報)で公開されている(例えば、宮城県医師会ホームページ)。そこで、2009年に発生した新型インフルエンザのデータを題材として統計モデルの構築をおこなう。

各保健所における1年間のインフルエンザ罹患データの変動を単純な時系列トレンドと地理的な近隣地区からの伝播の和としてモデル化をおこなう (Paul et al. (2008), Paul and Held(2011), Meyer and Held(2014))。

さらに、小池(2011)を参考に、異なる空間集積レベルであるデータの同時利用に関する検証をおこなう。データは、宮城県の各種人口データを用いる。より具体的には、統計情報研究開発センターから市販されている3次および4次メッシュデータ、基本単位区別集計データと政府統計の総合窓口で公開されている町丁・字等別集計データである。

結果の公表の最小地域単位である基本単位区別集計データは、公表項目が限られる。

一方、メッシュ統計は地域単位が広いものの、公表されるデータ項目が多いため、利便性は高い。したがって、メッシュデータがどの程度の利用可能性を見せるかが、大きな関心事である。

以上のデータを使って、面積按分法を用いた人口推計問題の検討をおこなう。任意の領域を指定したときに、その領域内の人口を面積按分によって推計する場合、上記の各種データで推計結果にどの程度の違いが出るのかを調べる。

さらに、従来は単純な面積按分による手法が良く用いられるが、補助データを用いることで推計精度がどの程度向上するかについても調査する。

4. 研究成果

時点 t における保健所管轄地域 i の罹患数 $Y_{i,t}$ は平均

$$\mu_{it} = e_i v_t + \lambda Y_{i,t-1} + \phi \sum_{j \neq i} w_{ji} Y_{j,t-1}$$

$$\log(v_t) = \alpha + \beta t$$

$$+ \gamma \sin(2\pi t / 52) + \delta \cos(2\pi t / 52)$$

であるとする (Paul et al. (2008), Paul and Held(2011), Meyer and Held(2014))。ここで、 e_i は地域 i の人口、 w_{ji} は隣接情報を表す。つまり、地域 i と地域 j が隣接していれば1、隣接していなければ0である。各保健所における1年間のインフルエンザ罹患数データの変動は、時系列トレンドと地理的な伝播の和として表現されている。

宮城県の8つの保健所管轄地域ごとに報告される罹患数の週報データ $Y_{i,t}$ が負の二項分布に従うと仮定してモデルパラメータの推定をおこなった。推定値は次のようになった：

$$\lambda = 0.64, \quad \phi = 0.05,$$

$$\alpha = 0, \quad \beta = 0.93, \quad \gamma = 4.92, \quad \delta = 2.54.$$

また、過分散を表すパラメータに対する95%信頼区間を求めたところ、[0.20, 0.28]となった。これにより、ポアソン分布で想定さ

れるよりも分散が大きくなるのが分かる。同時に、AICによって負の二項分布モデルとポアソンモデルを比較したところ、やはり負の二項分布モデルが支持された。

推定された負の二項分布モデルにおいて、Paul et al. (2008)に従い、時系列トレンドと地理的伝播との影響力の大きさを調べたところ、 $Y_{i,t}$ の変動のおよそ80%が地理的伝播によるとの結果が得られた。

上記のモデルでは、人口データとして保健所管轄区域レベルのものを使用している。しかし、同一管轄地域内であっても、人口密度は一定ではなく、当然、密度が大きい小地域の方がインフルエンザ罹患のリスクは高まると考えられる。

小池(2011)を参考にして、仙台市における任意の領域をバッファリングした場合、その圏内人口の推計精度について検証した。推計に使うデータは4次メッシュ、3次メッシュ、町丁・字等別集計である。また、推計法として、以下の3つの方法を比較した。

重心：推計領域に重心が含まれるメッシュの人口の総和

面積按分：メッシュの人口を推計領域に含まれる面積割合で按分する方法

基本単位区世帯数面積按分：人口と強い相関を持つ世帯数を補助データとして用いた按分法

結果として、仙台市におけるバッファリング圏内人口の推計には、統計データとして4次メッシュを、推計方法として基本単位区世帯数面積按分を用いる組み合わせが、最も推計精度は高かった。現在、空間相関を考慮した推計法についても検証中である。また、仙台市は政令指定都市であるため、5次メッシュデータが作成されている。これが利用可能な地域においては、これを積極的に用いた推計をすれば、さらに精度の良い結果が得られるであろう。

今後は、インフルエンザの点過程モデルに補助データを加える方向で更なる研究を継続する。現在のモデルでは、地域間の距離として、隣接しているか否かの情報しか用いていない。しかし、物理的に等距離であっても、人口の多い隣接地域からより強い影響を受けると考えられる。したがって、隣接地区との境界付近にどの程度の人口がいるのかという情報を小地域統計によってモデルに反映させることが必要である。また、人の往来も極めて重要な情報であるため、これについても併せて検討を要する。

引用文献

Nicola G. Best, Katja Ickstadt, Robert L. Wolpert. Spatial Poisson Regression for Health and Exposure Data Measured at Disparate Resolutions, *Journal of the American Statistical Association*, Vol. 95, 2000.

Carol A. Gotway, Linda J. Young. Combining Incompatible Spatial Data, *Journal of the American Statistical Association*, vol. 97, No. 458, pp. 632-648, 2002.

Michael Hohle. Additive-Multiplicative Regression Models for Spatio-Temporal Epidemics, *biometrical Journal*, 51(6), pp. 961-978, 2009.

Paul M, Held L, Predictive Assessment of a Non-Linear Random Effects Model for Multivariate Time Series of Infectious Disease Counts, *Statistics in Medicine*, Vol. 30, 1118-1136, 2011.

Paul M, Held L, Toschke A, Multivariate Modelling of Infectious Disease Surveillance Data, *Statistics in Medicine*, Vol. 27, 6250-6267, 2008.

稲葉寿. 感染症の数理モデル, 培風館, 2008.

小池司朗. 地域メッシュ統計の区画変遷に伴う時系列分析の可能性に関する一考察 測地系間・メッシュ階層間の比較から, 人口問題研究, 第67巻2号, 2011, 65-83.

Sebastian Meyer. Spatio-temporal infectious disease epidemiology based on point processes, Master's thesis, 2009.

Sebastian Meyer and Leonhard Held. Power-Law Models for Infectious Disease Spread, *The Annals of Applied Statistics*, Vol. 8, No. 3, 1612-1639, 2014.

宮城県医師会ホームページ

<http://www.miyagi.med.or.jp/> (2015年6月)

7 日閲覧)

政府統計の総合窓口
http://www.e-stat.go.jp/(2015 年 6 月 7 日閲覧)

高橋邦彦, 横山徹爾, 丹後敏郎. 疾病地図から疾病集積性へ, 保健医療科学, 第 57 号 2 巻, 2008, 86-92.

丹後俊郎, 横山徹爾, 高橋邦彦. 空間疫学への招待 疾病地図と疾病集積性を中心として, 朝倉書店, 2007.

5. 主な発表論文等
(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 0 件)

[学会発表](計 0 件)

[図書](計 0 件)

[産業財産権]
出願状況(計 0 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

取得状況(計 0 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
取得年月日:
国内外の別:

[その他]
ホームページ等

6. 研究組織

(1)研究代表者
坂口 隆之 (SAKAGUCHI, Takayuki)
山形大学・地域教育文化学部・講師
研究者番号: 10436496

(2)研究分担者
()

研究者番号:

(3)連携研究者
()

研究者番号: