

# 宿泊旅行統計を用いた観光の諸特性の分析

相馬 大<sup>1</sup>・兵藤 哲朗<sup>2</sup>

<sup>1</sup>学生会員 東京海洋大学大学院 修士 2 年 (〒135-8533 東京都江東区越中島 2-1-6)

E-mail: m195009@edu.kaiyodai.ac.jp

<sup>2</sup>正会員 東京海洋大学 流通情報工学科 教授 (〒135-8533 東京都江東区越中島 2-1-6)

E-mail: hyodo@kaiyodai.ac.jp

観光庁は 2007 年より宿泊旅行統計を実施している。同調査は宿泊施設に直接、宿泊者数やその国籍などを問う公的な統計であり、特に近年のインバウンドの激増時期をタイムリーに捉えた調査として高い価値を有する。調査の集計結果は、基本的に毎月ネット上でも公開されており、その情報を用いた実務や研究が積み重ねられてきたが、空間の解像度などに課題があった。本研究は、宿泊旅行統計の個票を用いて、施設の住所情報を geocoding し、3 次メッシュ (約 1km 四方) の空間情報を紐付けた『マスターデータ』を作成した。

これにより、高い空間精度を利した各種の分析が可能となり、かつ、公開されている集計表では把握できない、多重にわたるクロス分析も容易に実現できた。本研究では、その事例として、北陸新幹線の施設タイプ別の開業効果の推計や、12 年間の宿泊者国内重心位置の変遷など、マスターデータの特長を活かした分析結果を紹介している。

**Key Words:** *Travel and Lodging Statistics, geocoding, mesh data, inbound demand*

## 1. はじめに

観光庁は、その前身である総合政策局観光政策課時代から、わが国の観光政策に資する調査体系を確立してきた。その根幹の一つが、宿泊旅行統計であり、同調査は 2007 年 1 月より全国の宿泊施設を対象に開始された。2008 年の 10 月に観光庁が発足した後も調査は今に至るまで継続され、その間、インバウンド需要の激増の様子をつぶさに捉えてきた。

宿泊旅行統計には、日本人および現時点で 21 カ国 (その他を含む) の宿泊者数が補足できるし、一部の大規模施設については、日本人の居住都道府県の把握もなされている。しかしながら、ネット上で公開されている集計データでは、空間情報は都道府県までで、十分な解像度を確保できないし、多重のクロス分析も不可能である。空間解像度については、例えば柴田ら<sup>1)</sup>では都道府県施設名簿から、市町村への擬似的な分解を試みている。また、佐藤<sup>2), 4)</sup>、Sato<sup>3)</sup>では、限られた期間ではあるが、宿泊旅行統計の個票から 3 次メッシュの空間情報を紐付けした分析が試みられている。本研究は、後者の研究を分析時点で得られる全期間に延長し、かつ統合的なマスターデータの構築を試みることを目的としている。

本稿では、時間制約の関係で、基本的な集計分析を中心に紹介するが、莫大な Big Data と化した 12 年間、144 ヶ月の宿泊旅行統計を用いれば、極めて

多様な分析が可能となることが期待できる。

## 2. 宿泊旅行統計マスターデータの構築

2007 年第 1 四半期から調査が開始された宿泊旅行統計について、2018 年 12 月までの 12 年間、144 ヶ月のデータを統計法三十三条に基づいて、観光庁から個票を得た。個票は宿泊施設名簿と、アンケート調査結果の 2 種類の EXCEL ファイルが、47 都道府県別および月別 (2007 年から 2015 年第 1 四半期までは四半期別) に作成されており、ファイル数の合計は 7,332 個に及んでいた。効率的な分析を行うため、本研究ではこのデータを、一レコードが各宿泊施設の月単位となる、一つの csv ファイルに統合し、それを以降、『マスターデータ』と称することとする。12 年に及ぶデータなので、途中で調査対象国が追加されたり、国番号が変更されたりしているため、統合作業は全てプログラムを作成し、作業の経過が明確に残るようにした。

さらに、構築した統合データについて、geocoding を施した。用いたのは、R の library である、ggmap の "geocoding" というコマンドである。半角や全角数字が混在する住所テキストであったが、"gocoding" を用いることで、37,402 軒の施設に対して、その成功率は、 $(37402-19)/37402 \times 100 = 99.95\%$  と極めて効率的な作業となった。

表-1 マスターデータの変数一覧

1 年	32 問9タイ	63 問9静岡
2 月	33 問9マレーシア	64 問9愛知
3 自治体JIS	34 問9インド	65 問9三重
4 3次meshコード	35 問9オーストラリア	66 問9滋賀
5 FORM形式	36 問9インドネシア	67 問9京都
6 回収	37 問9ベトナム	68 問9大阪
7 調査票様式 (=5.FORM)	38 問9フィリピン	69 問9兵庫
8 問3施設タイプ	39 問9イタリア	70 問9奈良
9 問4A客室数	40 問9スペイン	71 問9和歌山
10 問4B収容人数	41 問9その他	72 問9鳥取
11 問5従業者数	42 問9北海道	73 問9島根
12 問6A観光割合	43 問9青森	74 問9岡山
13 問6B業務割合	44 問9岩手	75 問9広島
14 問7A1宿泊者数延べ人数	45 問9宮城	76 問9山口
15 問7A2宿泊者数実人数	46 問9秋田	77 問9徳島
16 問7A3外国人延べ人数	47 問9山形	78 問9香川
17 問7A4外国人実人数	48 問9福島	79 問9愛媛
18 問7A5利用客室数	49 問9茨城	80 問9高知
19 問8県内	50 問9栃木	81 問9福岡
20 問8県外	51 問9群馬	82 問9佐賀
21 問9韓国	52 問9埼玉	83 問9長崎
22 問9中国	53 問9千葉	84 問9熊本
23 問9香港	54 問9東京	85 問9大分
24 問9台湾	55 問9神奈川	86 問9宮崎
25 問9アメリカ	56 問9新潟	87 問9鹿児島
26 問9カナダ	57 問9富山	88 問9沖縄
27 問9イギリス	58 問9石川	89 問9国外
28 問9ドイツ	59 問9福井	90 外れ値FLG
29 問9フランス	60 問9山梨	91 weight
30 問9ロシア	61 問9長野	92 従業者区分
31 問9シンガポール	62 問9岐阜	

構築したマスターデータには、表-1 に示す変数が含まれ、144ヶ月の全レコード数は、1,480,089におよび、csvファイルのファイルサイズは、約380Mバイトである。

### 3. 宿泊旅行統計の基礎集計

#### (1) 調査票様式の構成率

宿泊旅行統計は、母集団である宿泊施設を従業員の規模別に分類し、その規模別に抽出率を定めている(表-2)。なお、表中に「悉皆」とあるが、回答率が100%ではないので、表中の抽出率は上限値である。

表-2 調査票種別の概略

調査票種別	従業者区分	施設授業者数	母集団数	抽出率
3号様式	5	100人以上	1,019	悉皆
2号様式	4	30~99人	4,190	
	3	10~29人	7,215	
1号様式	2	5~9人	6,875	1/3
	1	0~4人	35,568	1/9

さらに、表の様式別に調査票が異なっている。表-1における変数との対応をまとめると、以下のとおりとなる。

1号様式：～問8まで

2号様式：～問9の国籍別宿泊者延べ人数まで

3号様式：～問9の居住地別宿泊者延べ人数

様式別のサンプル数(図-1)と、日本人と外国人を合わせた延べ宿泊者数(図-2)を見ると、2010年4月から、調査対象数が5割程度増加し、精度の向上が図られてきたことが確認できる。なお、後者の宿泊者数は、個々のサンプルについて、表-1の「14: 問7A1 宿泊者数延べ人数」に「91: weight(もしくは乗率, 拡大係数)」を乗じた値を用いている。また、2007年の開始から2010年3月までの様式2は2010年4月以降の様式3に、様式1は様式2に対応しており、現在の様式1は2010年4月以降導入されてきたことに留意する必要がある。なお、図では2007年度はNAになっているが、これは今回入手したデータにおいて様式の情報がなかったためである。図-1と図-2の比較から、サンプル数においても、宿泊者数においても2010年4月以降の様式2が半数程度を占めており、宿泊旅行統計において重要な調査対象であることが理解できる。

#### (2) 施設タイプの構成率

表-1の「8: 問3 施設タイプ」別に集計した結果を図-3と図-4に示す。サンプル数では旅館の割合が高いものの、宿泊者数では近年特にビジネスホテルのシェアが高まっており、同施設の大規模化が想像される。

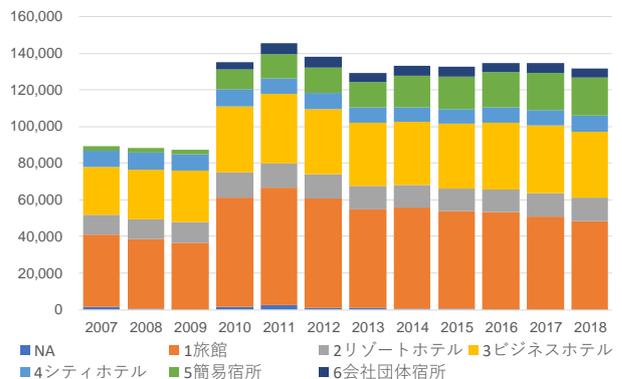


図-3 施設タイプ別の年別サンプル数

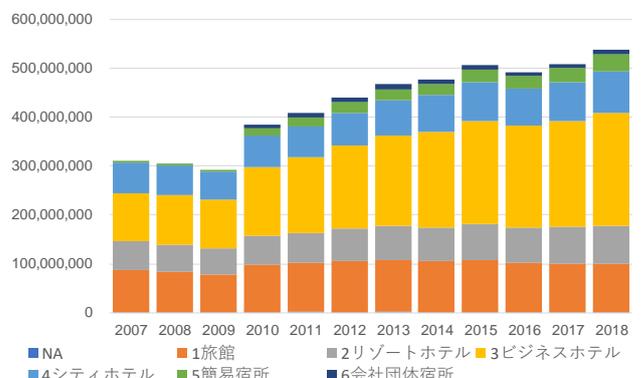


図-4 施設タイプ別の年別延べ宿泊者数

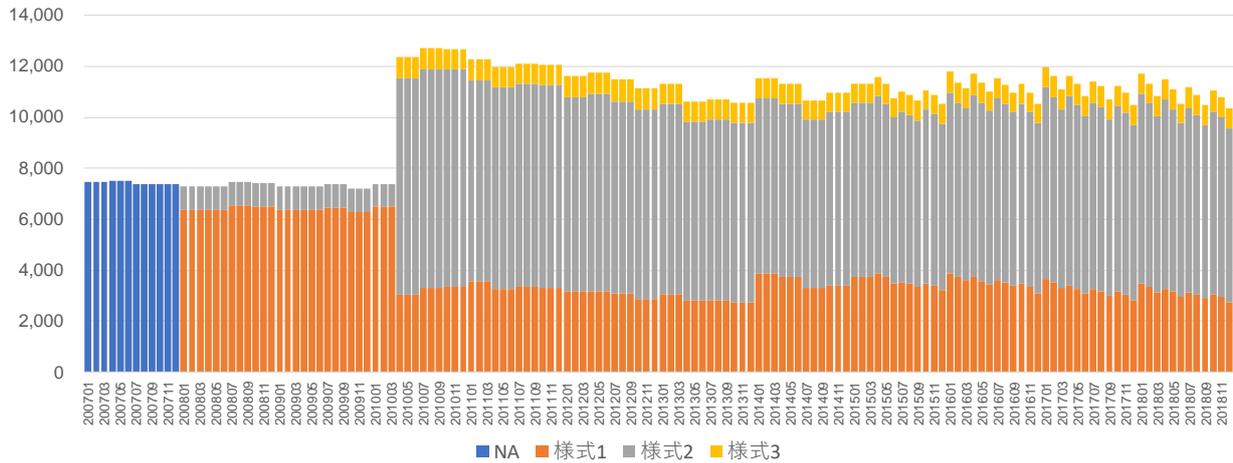


図-1 様式別の月別サンプル数

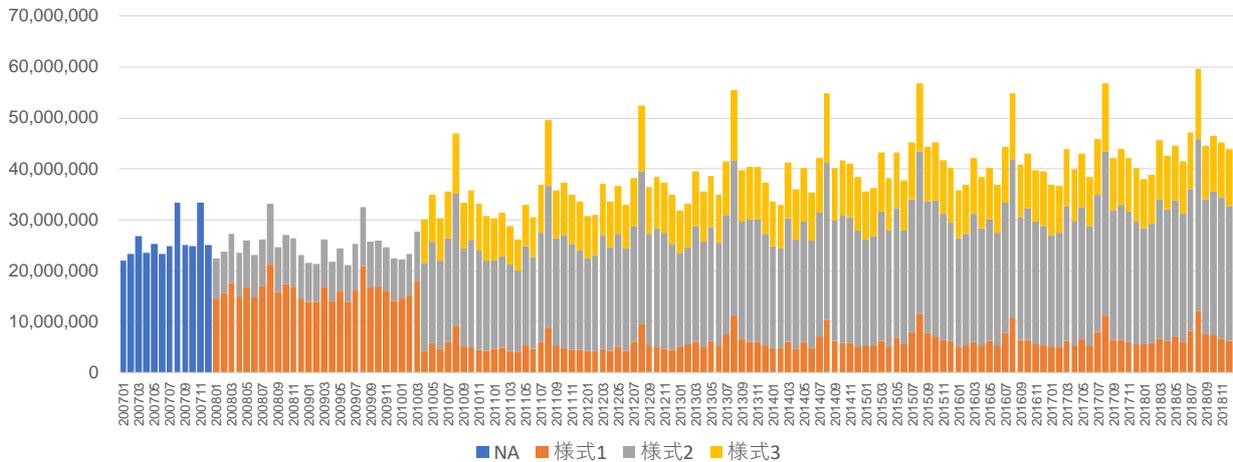


図-2 様式別の月別延べ宿泊者数

(3) 宿泊者数に関する分析

図-5 に日本人と外国人の月間宿泊者数の推移を、図-6 に外国人の月間宿泊者数割合推移を示す。日本人宿泊者が大半を占めているが、年次を追うごとに外国人宿泊者が増加していることが分かる。2011年の東日本大震災のときに一時急減するが、減少した需要は回復し、近年では最大で全体の20%を占める月も見られる。これは震災の影響で日本が安全な国ではないと一時的に見られたが、年を追うごとに復興が進んだことや、2020年の東京オリンピックを控えていることによると考えられる。

ここで施設タイプ別の宿泊者推移を確認していく。図-7 に施設数の推移を、図-8 から図-13 に施設タイプ別の日本人と外国人の月間宿泊者数推移を示す。図-7 より項目「0:不明」「9:不明」などが見られたため、施設タイプ別宿泊者数を示す際にはこれらのデータを除いて集計するものとする。近年では旅館が減少している一方で簡易宿所が増加傾向にある。これは訪日外国人増加に伴い、山小屋などの観光地宿泊施設やカプセルホテルなどの手軽な宿泊施設が建設されたためだと考えられる。

図-8 から図-13 より宿泊者数が最多なのは施設タイプ3のビジネスホテルであることが分かる。施設

数としては施設タイプ1の旅館が最多であるが、ビジネスホテルの宿泊者が多い理由としては、1泊当たりの値段が関係していると考えられる。旅館には食事がついている分、ビジネスホテルよりも高くなりがちである。そのため、仕事などで利用する際や、多少の旅行ではビジネスホテルを利用する人が多くなる。また、施設タイプ1, 2, 3, 5, 6では日本人宿泊者の割合が高いが、施設タイプ4は外国人宿泊者の割合が他の施設タイプと比べて高いことがわかる。シティホテルはビジネスホテルよりも宿泊費が高いが、お客様をサポートするスタッフが常駐していることが多く、高級レストランやバーなどのテナントが入っていることや都心部の近く、眺望がいい場所に作られている場合が多い。宿泊する以外にも快適に過ごすサービスが整えられているため、見知らぬ土地で過ごす不安を解消したい外国人利用者が多くなったのだと考えられる。図-14 に国別外国人の月間宿泊者数推移を示す。中国、台湾、韓国、香港などのアジア圏内の国からの宿泊者が多く、その他だとアメリカからの宿泊者が多い傾向となっている。次の章では宿泊者数が多い韓国、中国、香港、台湾、アメリカに着目し分析を行う。

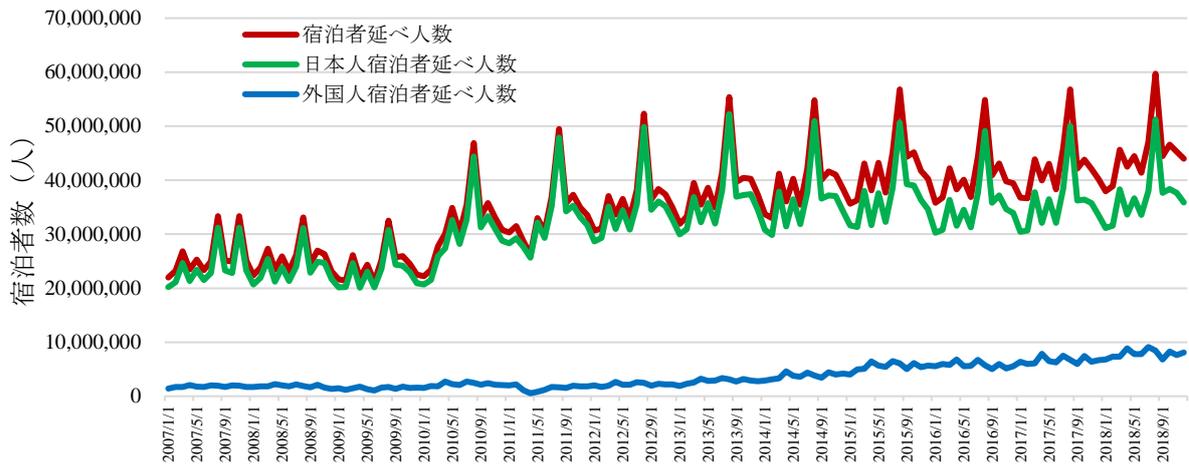


図-5 日本人と外国人の月間宿泊者数推移

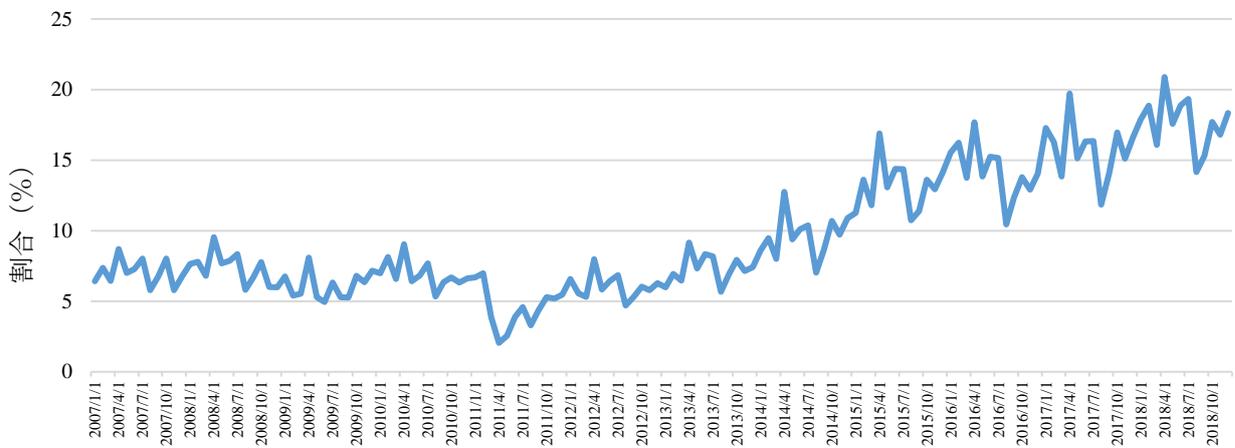


図-6 外国人の月間宿泊者数割合推移

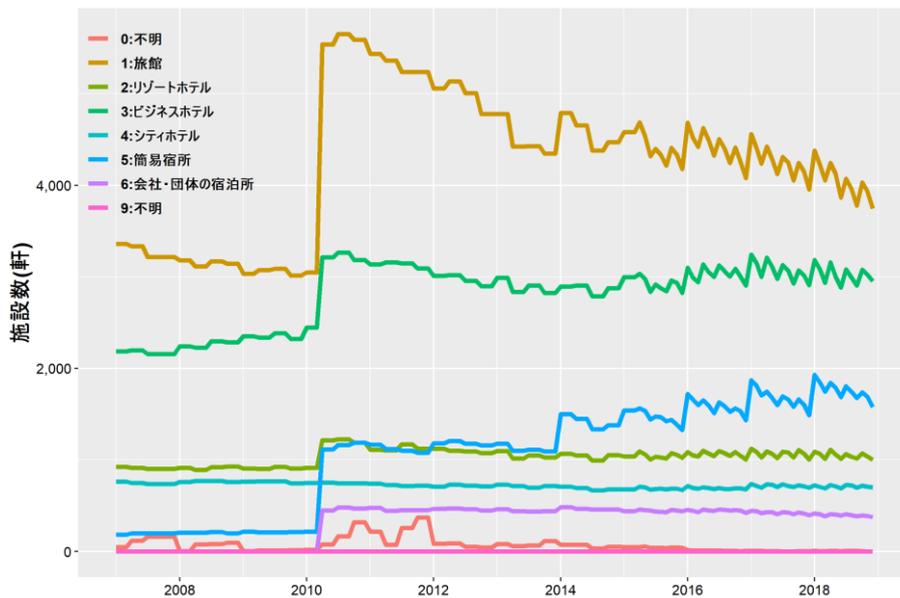


図-7 施設数推移

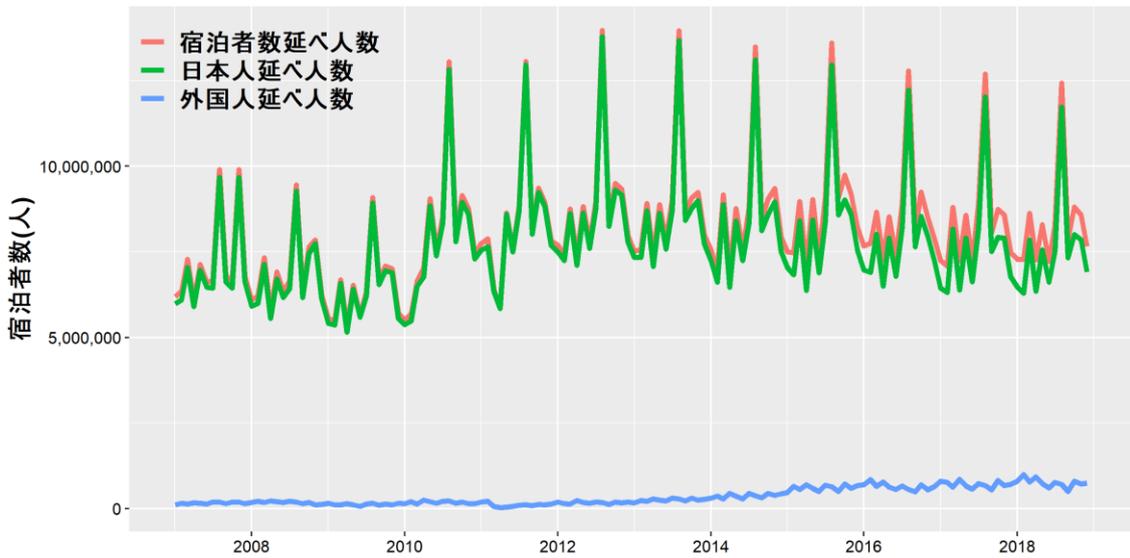


図-8 施設タイプ 1 (旅館) の宿泊者数推移

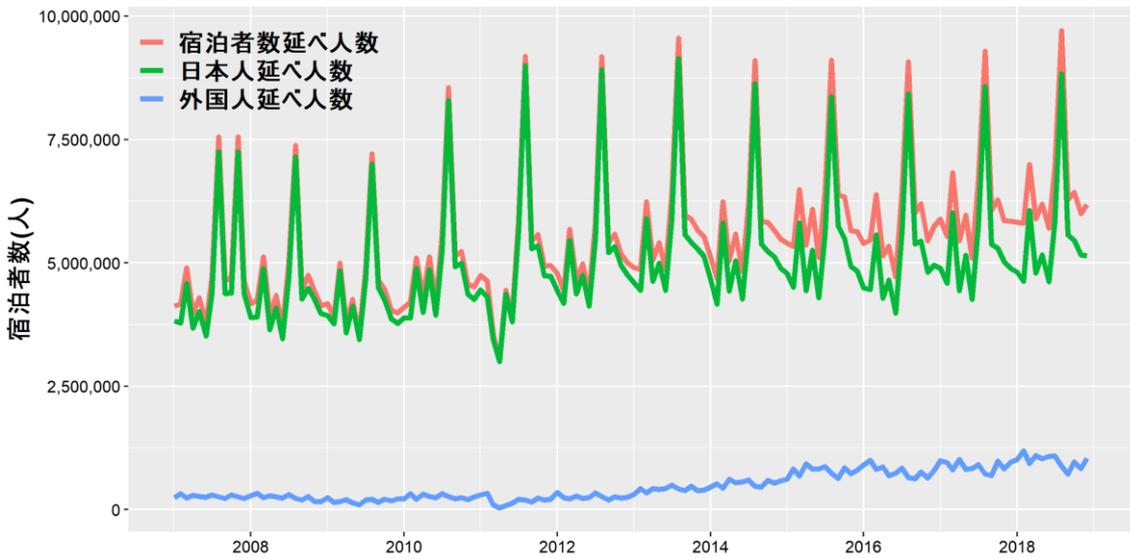


図-9 施設タイプ 2 (リゾートホテル) の宿泊者数推移

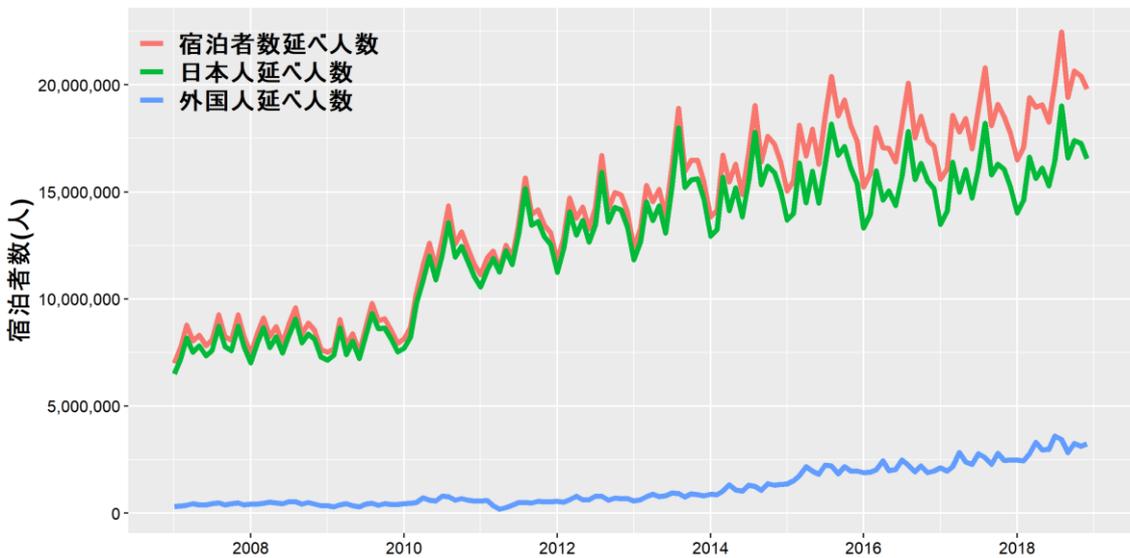


図-10 施設タイプ 3 (ビジネスホテル) の宿泊者数推移

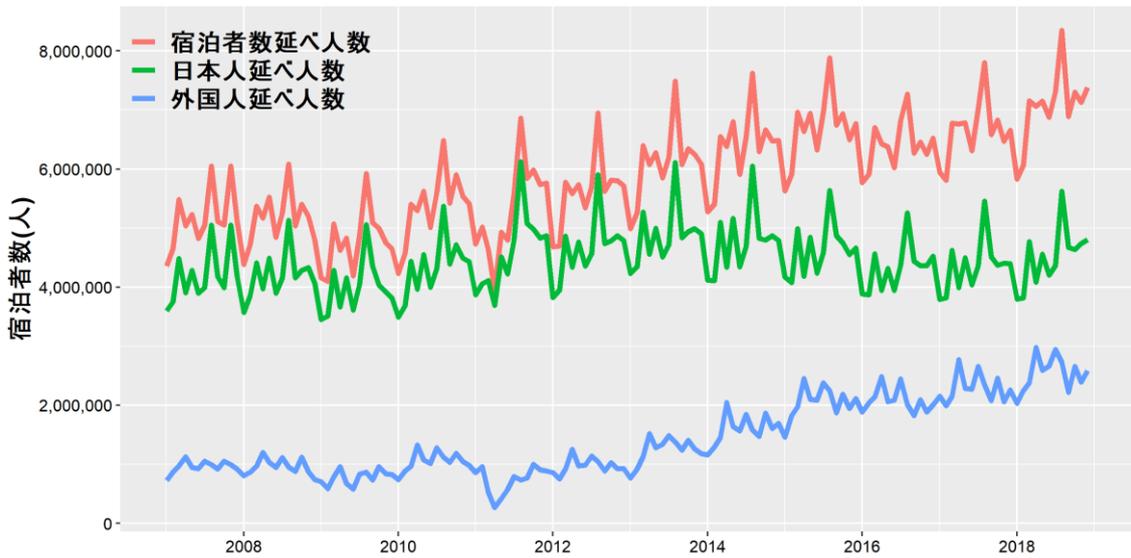


図-11 施設タイプ 4 (シティホテル) の宿泊者数推移

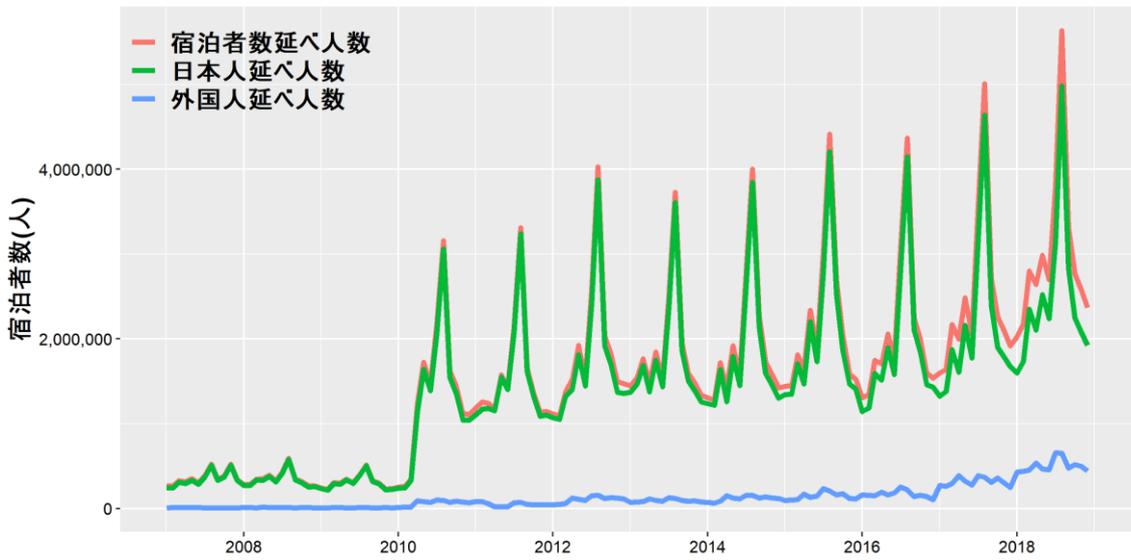


図-12 施設タイプ 5 (簡易宿所) の宿泊者数推移

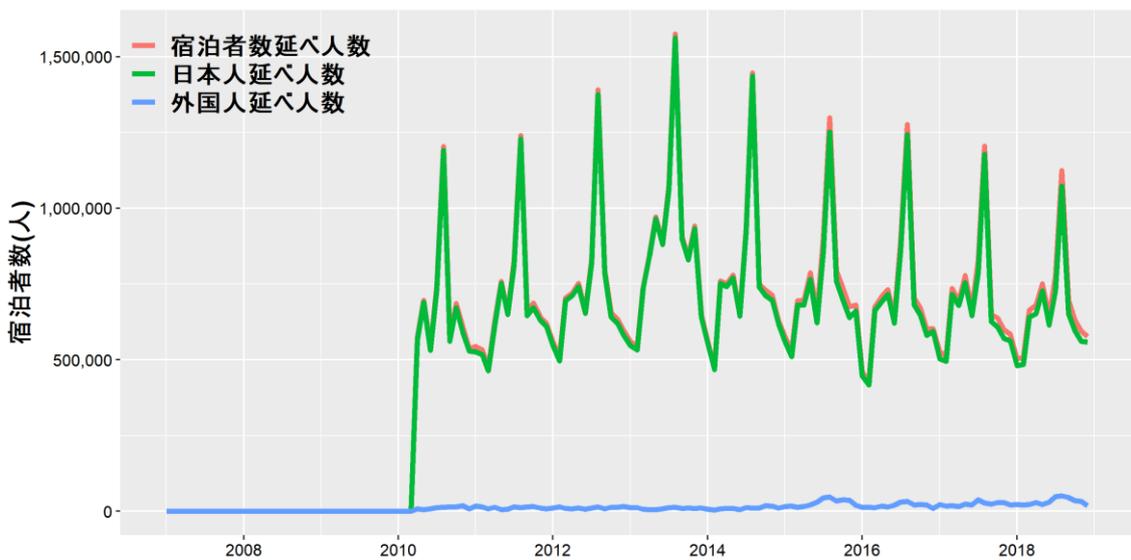


図-13 施設タイプ 6 (会社・団体の宿泊所) の宿泊者数推移

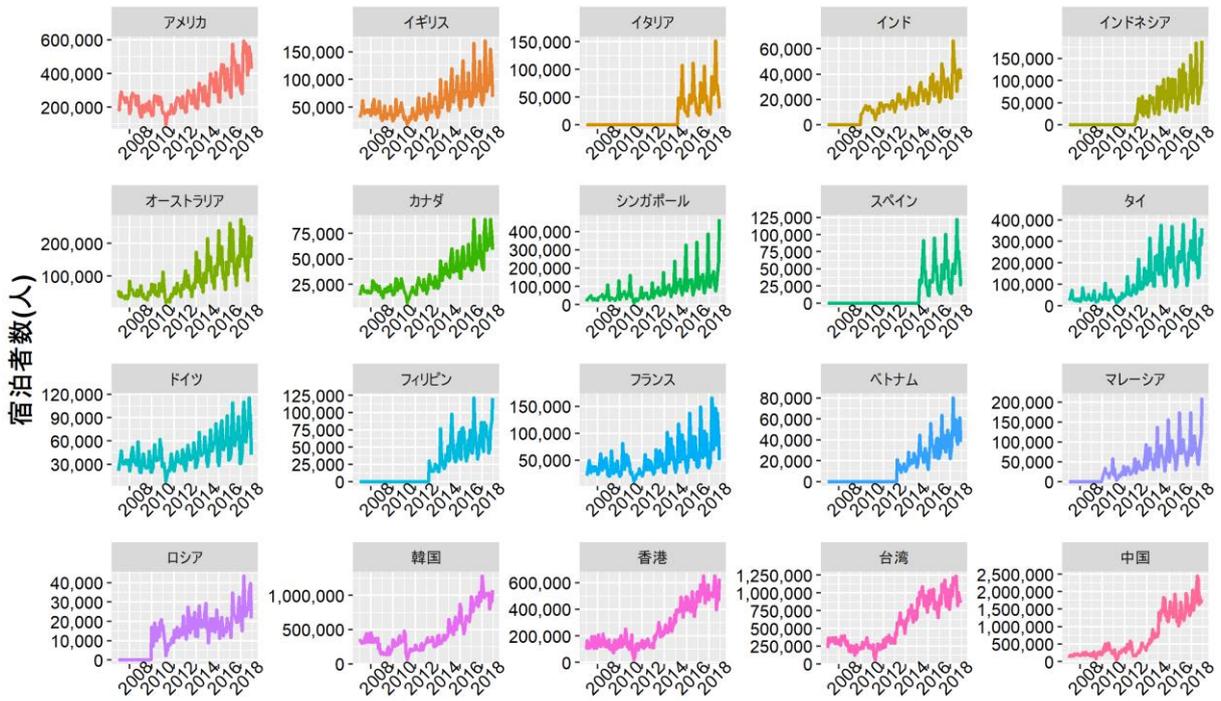


図-14 国別外国人宿泊者数推移

#### 4. 個別トピックに関する分析

##### (1) 北陸新幹線開業効果の分析

調査票の3号様式の間9では、宿泊者の居住地を都道府県および国外の区分で質問している。これを用いれば、例えば交通施設整備の前後の宿泊客数の変化などを確認することができる。ここでは2015年3月の北陸新幹線の開業効果を確認してみる。ただし、間9の居住地項目は3号様式施設に限られるため、図-2から見て取れるように、全体の3割程度のサンプル率であることには留意する必要がある。

分析は、宿泊施設を富山県と石川県に、そして関東からの訪問客が増えたことが知られていることから、居住地には、埼玉県・千葉県・東京都・神奈川県を選んだ。なお、これだけであればネットで公開されている集計表でも計算できるので、ここではマスターデータの特長を活かして、施設タイプを「旅館」と「ビジネスホテル+シティホテル」に区分した集計を行った。

図-15と図-16より、旅館もホテルも新幹線開業直後から多くの首都圏からの宿泊客を迎えていることが分かるが、ホテルにおいてその傾向が顕著であり、2016年以降、元の2014年レベルに戻っていることが意外である。しかし旅館ではそれほどまで2014年の値に近づいているわけではない。旅行者の属性や移動目的と相関の高い施設タイプ別の分析から、開業効果に差異があることを見出すことができたと考える。

##### (2) 様々な宿泊属性別の重心位置の変遷

今回のマスターデータの特長の一つはgeocodingにより、1km単位の3次メッシュ情報を用いることができ

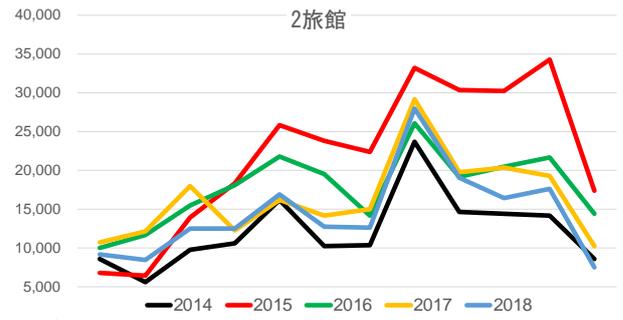


図-15 「富山・石川県←首都圏4都県」の月別延べ宿泊者数

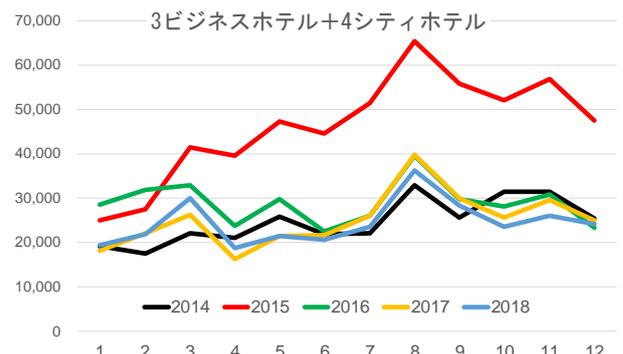


図-16 「富山・石川県←首都圏4都県」の月別延べ宿泊者数

ることにある。ここでは3次メッシュの緯度経度情報を用いて、延べ宿泊者数の年別の重心を、「日本人+外国人(ALL)」「外国人」「Korea」「China」「HongKong」「Taiwan」「USA」の属性別に算出し、日本地図上にプロットした(図-17)。紙面の都合上、偶数年のみの掲載となるが、判明したことは以下のとおりである。

- ・2008年時点では、台湾と香港の重心は長野県北部、

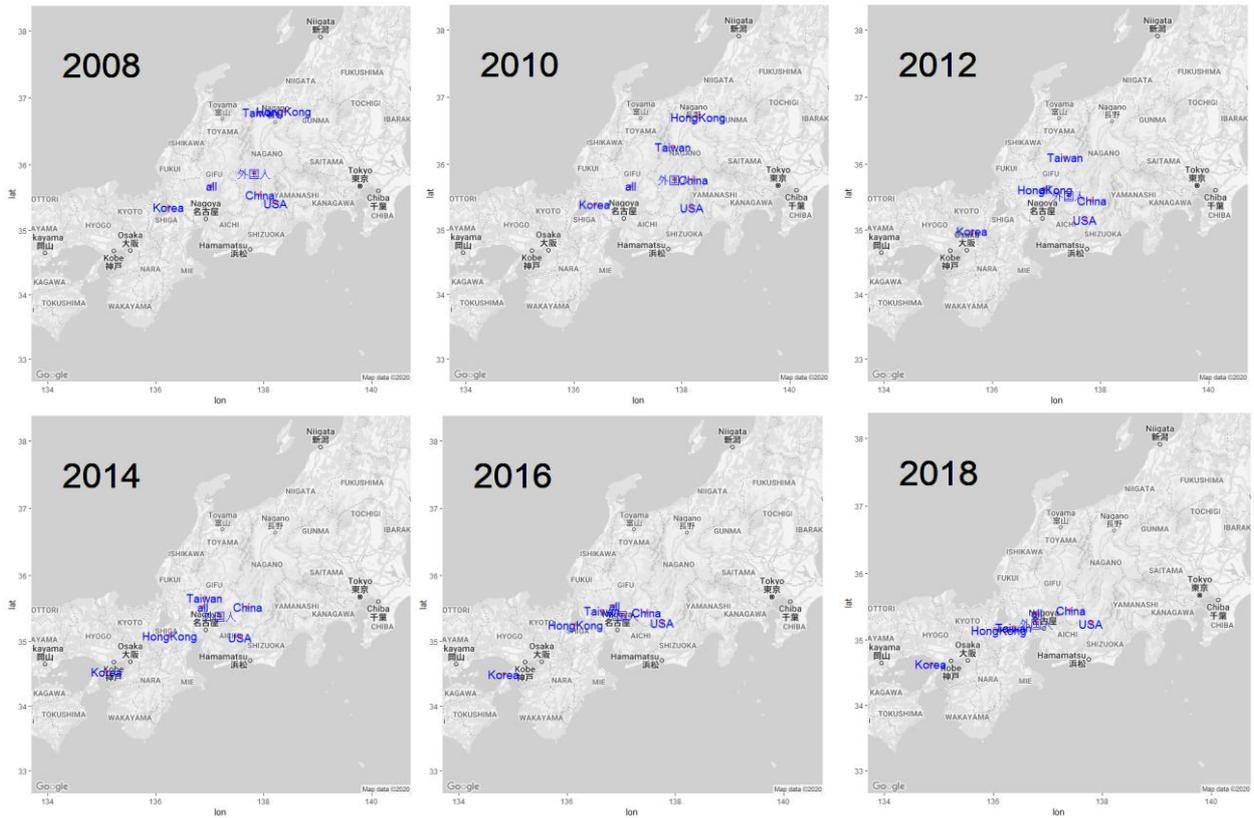


図-17 宿泊者数の属性別重心位置の変遷

中国とアメリカの重心は長野県南部，韓国は滋賀県にあった。

- ・2010年以降，台湾・香港・韓国の重心は徐々に南西部に移動して，台湾・香港は滋賀県周辺，韓国は淡路島周辺にとどまっている。
- ・中国とアメリカの重心は比較的安定した場所に位置している。

これらの原因としては，台湾・香港・韓国の沖縄県への訪問者が激増したことが想像された。具体的に，この5カ国の年別沖縄県延べ宿泊率を確認してみると（図-18）確かに，中国・アメリカに比して沖縄県を訪れる割合が増加していることが分かる。

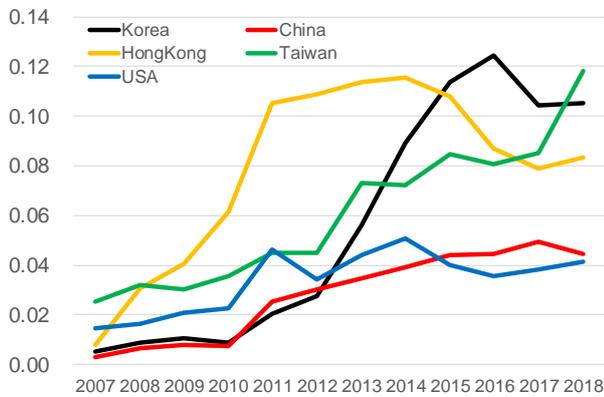


図-18 5カ国の年別沖縄県宿泊率（延べ宿泊数）

## (2) メッシュにおける延べ宿泊者数と外国人宿泊者数に関する回帰分析

本項では，佐藤<sup>2),4)</sup>が行った地域メッシュ統計を用いた回帰分析を，サンプル数を144ヶ月に拡大して検討する。この分析では年月・3次メッシュごとに延べ宿泊者数，国籍別延べ外国人宿泊者数を集計し，メッシュ $c$ における延べ宿泊者数 $X(c)$ を説明変数，同じメッシュでの国籍別外国人延べ宿泊者数 $Y(c)$ として，式(1)の単回帰を推定する。

$$\log_{10} Y(c) = \alpha \log_{10} X(c) + \beta \quad (1)$$

$Y(c)$  : メッシュ $c$ における国籍別外国人延べ宿泊者数，

$X(c)$  : メッシュ $c$ における延べ宿泊者数，

$\alpha$  : 回帰係数， $\beta$  : 切片

分析対象は，前節と同じ5カ国で，図-19に調整済み決定係数の推移を，図-20に回帰係数の推移を示す。自由度調整済み決定係数はどの国も0.3近辺であり，年次が経つにつれて上昇し，近年では韓国を除いて0.4を超えるに至る。このことから，メッシュ $c$ における延べ宿泊者数 $X(c)$ は国籍別外国人延べ宿泊者数 $Y(c)$ に対して，十分な説明能力を持たないと言えるが，年次が経つにつれて説明力は向上していると考えられる。また，回帰係数についても同様で，年次が経つごとに係数も上昇している。これらは，訪日外国人宿泊者数が増加していることや，サンプル数が伸びていることや調査対象旅館が増加しているためと考えられる。

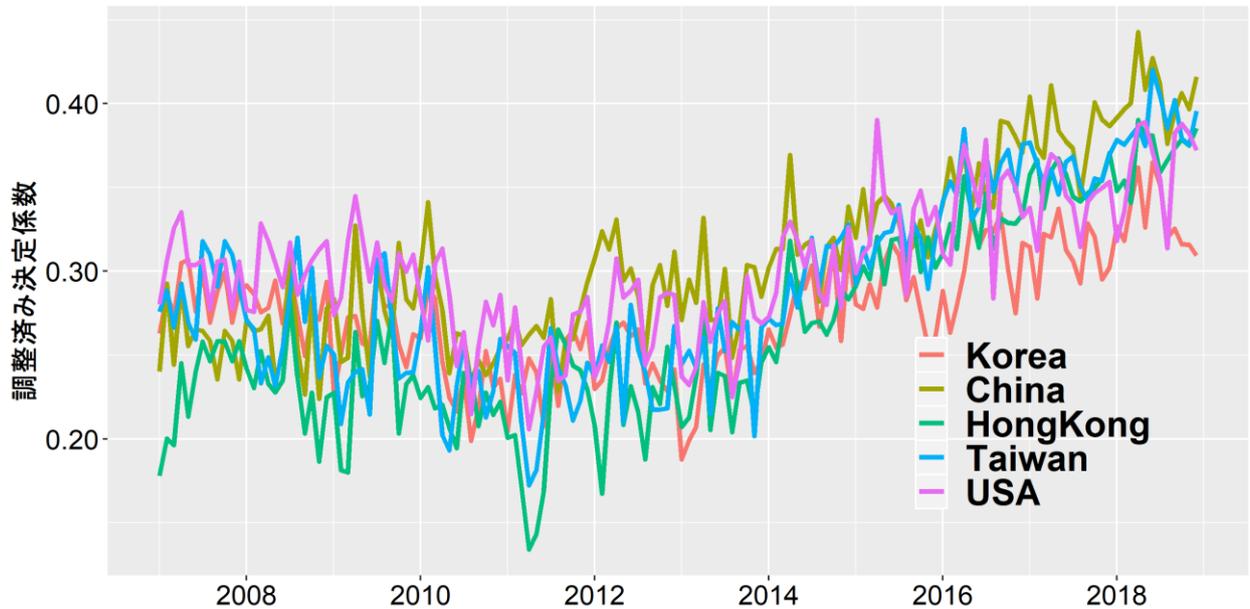


図-19 調整済み決定係数推移

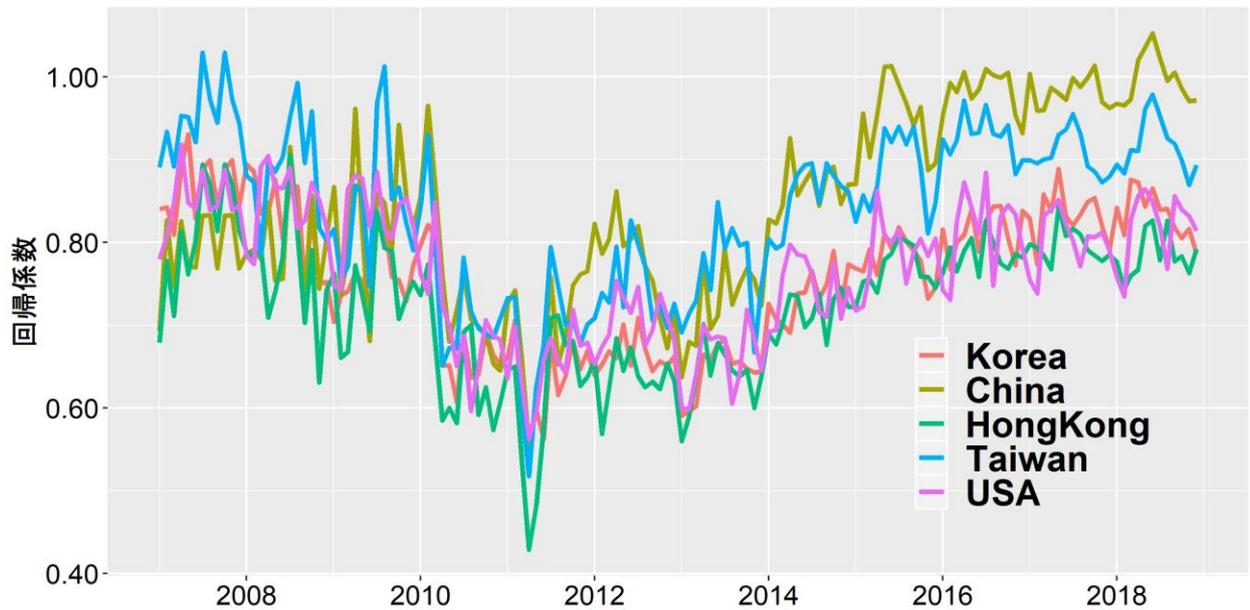


図-20 回帰係数推移

## 5. おわりに

本研究では、2007年1月から2018年の12月までの144か月分かつ3次メッシュや調査様式別に、国籍別・都道府県別の宿泊者数が分析可能なマスターデータを構築し、各種の集計分析を行った。宿泊者の大半は日本人であるが、年月を追うごとに外国人宿泊者は伸びており、近年では全体の2割を占めるまでになった。また、宿泊施設タイプ別に日本人宿泊者数と外国人宿泊者数を分析してみると、日本人ではビジネスホテルが多く、外国人宿泊者ではシティホテルが多いことが分かった。

また北陸新幹線の開業効果については、開業当時は宿泊者数は大きく増加したが、近年では開業前と

同程度まで落ち着いていることが分かった。国籍別宿泊者数の重心位置については年月を追うごとに沖縄県の宿泊者数が増加しており、南西に推移していることが分かった。

最後に回帰分析によるパラメータ推移については、回帰係数が年月を追うごとに大きくなっており、各国宿泊者数の、宿泊者総数に対する弾力性が高まっていることも判明した。

今後の課題としては、3次メッシュごとに宿泊者数を求め、どのような地点で宿泊者が多くなるかを調査することや、宿泊施設ごとの稼働率、従業員区分を用いた分析を行い、訪日外国人の特性を検討していくことが必要である。

**謝辞**：末筆ではあるが、宿泊旅行統計データを得るに当たり、観光庁観光戦略課観光統計調査室の方々には多大なご協力を頂いた。ここに記して謝意を表す次第である。

**参考文献**

- 1) 柴田優作, 日比野直彦, 森地茂：宿泊旅行統計を用いた訪日外国人の市町村別宿泊実態の把握, 土木学会論文集 D3, Vol.75, No.5, pp.461-474, 2019
- 2) 佐藤彰洋：メッシュ統計, 共立出版, 2019
- 3) Sato,A.-H.: "Microdata analysis of the accommodation survey in Japanese tourism statistics", 2015 IEEE International Conference on Big Data, Oct. 29 2015 – Nov. 1, pp.2700-2708
- 4) 佐藤彰洋：政府統計を用いた国内観光状況の網羅的分析, 情報処理学会第 78 回全国大会, pp.1-265-1-266, 2019

**(Received March 6, 2020)**

**A STUDY ON CHARACTERISTICS OF OVERNIGHT TOURIST DEMAND  
BY TRAVEL AND LODGING STATISTICS IN JAPAN**

Dai SOMA, Tetsuro HYODO