

2019年度 統計データ分析コンペティション

特別賞（大学生・一般の部）

外国人人口と市区町村の特性との関係性

西尾 春香（関西学院大学経済学部）

論文の概要

外国人人口と自治体の特性との関係について分析するため、市区町村別の統計指標を用いた重回帰分析を行い、人口密度や製造業等との相関が高いという結果を得た。また、決定木分析、ランダムフォレスト等の追加分析を行い、都市化、働き手不足、特定産業等が外国人人口に影響していることを指摘している。

論文審査会コメント

外国人人口について様々な探索的なデータ分析を行い、機械学習による政策分析の可能性を示す論文として評価された。論文としては47件に対して29変数の投入の問題など、再考の余地があるが解釈も明確である。

外国人人口と市区町村の特性との関係性

西尾 春香

関西学院大学経済学部

1. 研究の目的

近年、日本に住む外国人に対して大きな関心が集まっている^{(1) (2)}。その背景として、1980年代後半から外国人人口の増加が顕著になり、リーマンショックや東日本大震災の影響で一時的に減少する時期もあるものの、最近も依然として増加傾向にある。さらに、2015年以降は対前年増減率においても増加の割合が大きくなっており、総務省が発表した2019年1月1日時点での住民基本台帳に基づく人口、人口動態及び世帯数調査によると、外国人人口は前年より約17万人増えて過去最多、さらに全国の人口総数のうち外国人住民の占める割合が初めて2%を超えた。一方で、日本人人口は前年より約43万人少なく、10年連続の減少となった。このように、日本人人口が減少を続けることで不足する労働力を補うためにも年々増加する外国人は欠かせない存在となっている。加えて、2019年4月1日より人手不足が深刻な産業分野において在留資格「特定技能」が新設され、一定の専門性・技能を有する新たな外国人材の受入れが可能となった。特に、熟練者を想定した特定技能2号では定期的な審査を経て、家族の帯同が認められる長期の滞在が可能になる。これらの点を踏まえると、在留外国人の増加傾向は今後も続くと考えられる。それに伴い、受け入れの各自治体では自然災害時の多言語対応や子弟の就学支援など、異なる文化的背景を持つ外国人が生活するための環境整備に早急に取り組む必要がある⁽³⁾。

本研究では外国人人口と自治体が持つ特性との関係性について計量分析を行う。各自治体は人口、経済、行政、教育、社会保障など多面的な特性を持っているが、そこに住む外国人の数とそれらの特性に何か結びつきがあるのか、加えて、もし関連があるならばそれはどの特性で、どの程度の関係性の強さなのかを調べることは、現在の日本の社会構造を理解する上で意味があると考えられる。これまでに行われた外国人人口に関する先行研究は、外国人の新規流入と国内移動⁽⁴⁾、市区町村における外国人の社会増加⁽⁵⁾、外国人女性の出生数と出生率^{(6) (7)}、外国人の居住地選択⁽⁸⁾、外国人労働者導入の日本人労働者・企業への影響⁽⁹⁾、都道府県別の外国人の自然動態⁽¹⁰⁾など多岐に渡るものの、外国人人口と市区町村の特性との関連性については十分に調べられていると言えず、本論文では教育用標準データセットの市区町村データを使った統計分析を通してこの点について考察を深めていきたい。

2. 分析の方法と手順

本研究の分析においては、外国人人口（10万人あたり）を目的変数、自治体の多様な特性を説明変数として、両者の間に存在する関係性をいくつかの統計手法を用いて推測する。外国人人口と自治体特性の関連を分析するための簡潔な手法としては、個々の特性と外国人人口の相関係数を個別に計算することが考えられるが、この方法では説明変数間の相関関係を無視することになる。そこで、本研究では重回帰分析を行うことで、外国人人口とすべての特性の関係性を同時に推定する。加えて、分析においては自治体の多面的な特性を十分記述するために多くの説明変数を用いるが（表1参照）、それらすべてが実際に外国人人口と関係があるとは限らないため、結果としてパラメータ数が多くなり過ぎてデータへの過適合を招く恐れがある。そこで、Best Subset Selectionを用いて説明変数の中から目的変数をよく説明すると考えられる部分集合を

選び出す。

重回帰分析は目的変数と説明変数の線形関係に基づくが、近年、この線型性を仮定しない機械学習の統計手法が自然科学、社会科学問わずに様々な分野で広く用いられている⁽¹¹⁾。そこで、データを段階的に分割する木構造に基づく決定木回帰、さらに複数の木を組み合わせることで予測精度を向上させるランダムフォレストも追加して分析を行う。

最後に、これら4つの分析手法の推定結果を得たあとで10-foldクロスバリデーションを行い、どの分析手法がもっとも妥当であるかを比較する。これらの統計分析にあたっては、統計解析向けのプログラミング言語であるR言語を用いた。Best Subset Selection、決定木、ランダムフォレストの推定においてはライブラリー“leaps”、“tree”、“RandomForest”をそれぞれ用いた。

3. 使用するデータについて

表1は、本論文の分析に用いる変数を一覧にしたものである。変数を社会・人口統計体系の中の項目定義のサイト(https://www.e-stat.go.jp/koumoku/koumoku_teigi/A)を参考に分野ごとに記載した。分析前のデータ処理として、分布が大きく歪んでいる非負データには対数変換を施している(0を含んでいる場合は1を加えたものを対数変換している)。一部の変数は総務省統計局「都道府県・市区町村のすがた(社会・人口統計体系)」⁽¹²⁾から追加的に取り出しており、表1では追加した変数の枠内に色付けをしている。具体的には、未婚者割合(15歳以上人口)、課税対象所得(納税義務者1人当たり)、学歴割合(大卒)、刑法犯認知件数であり、それぞれの変数の収録年度は順に、2015年、2017年、2010年、2008年である。経済基盤の従業者割合に関しては、外国人労働者数が比較的多い産業を選択した⁽¹³⁾。

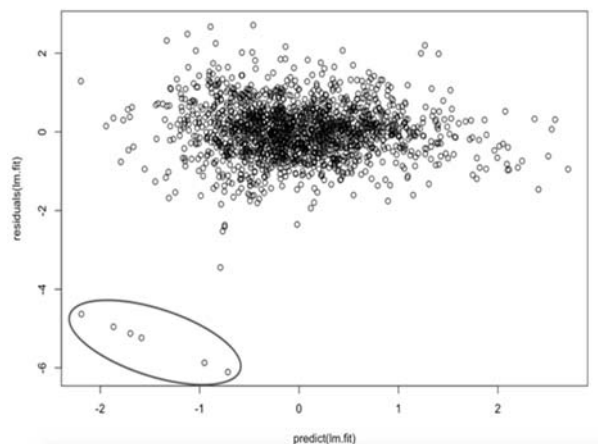
また、総人口がゼロの市区町村と、福島県檜葉町(R07542)、富岡町(R07543)、大熊町(R07545)、双葉町(R07546)、浪江町(R07547)、葛尾村(R07548)、飯館村(R07564)は「農林業センサス」を出典とする項目について調査が実施されなかったため取り除いている⁽¹⁴⁾。最終的に、これらの変数をそれぞれ平均0、分散1になるように標準化したものを分析に使用した。

4. 分析結果

4.1 重回帰分析

最初に、すべての説明変数を使った重回帰モデルの推定を行い、残差プロットを調べたところ(図1)外れ値と考えられるサンプル(円で囲んでいる部分)が6箇所ほど観測された。これらが推定結果に大きく影響する可能性があるため、データから取り除いて分析を行うことにした。比較の観点から、以下述べる他の統計手法に対しても同じデータを用いることにする。表2が上の処理を行ったデータを用いた重回帰モデルの推定結果であり、様々な分野の項目の係数が有意に推定されていることがわかる。例えば、人口密度(人口・世帯)、製造業従業者割合(経済基盤)、課税対象所得

図1：残差プロット

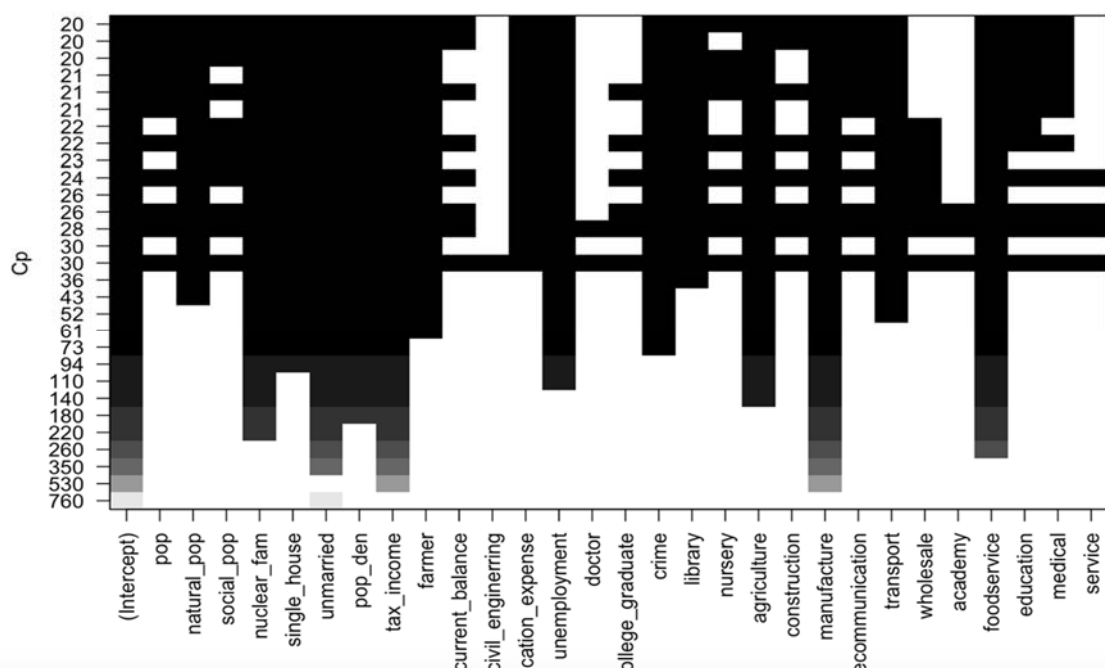


(行政)、図書館数(文化・スポーツ)などは正の係数を持ち、自然人口増減率(人口・世帯)や完全失業率(労働)は負の関係性を示している。分析前のデータ処理ですべての変数を標準化しているため、回帰係数

表 1：分析に用いる変数

分野	変数名	計算式	単位
人口・世帯	foreigner 【外国人人口(10万人あたり)】	$\log(\text{外国人人口} / \text{総人口} \times 100000 + 1)$	人
	pop 【総人口】	$\log(\text{総人口})$	人
	natural_pop 【自然人口増減率】	$(\text{出生数} - \text{死亡数}) / \text{総人口}$	%
	social_pop 【社会的人口増減率】	$(\text{転入者数} - \text{転出者数}) / \text{総人口}$	%
	nuclear_fam 【核家族世帯割合】	$\text{核家族世帯数} / \text{一般世帯数} \times 100$	%
	single_house 【単独世帯割合】	$\text{単独世帯数} / \text{一般世帯数} \times 100$	%
	pop_den 【人口密度】	$\log(\text{総人口} / \text{可住地面積})$	人/ha
	unmarried 【未婚者割合(15歳以上人口)】	未婚者割合(15歳以上人口)	%
経済基盤	agriculture 【従業者割合(農業、林業)】	$\text{従業者数(農業、林業)} / \text{従業者数} \times 100$	%
	construction 【従業者割合(建設業)】	$\text{従業者数(建設業)} / \text{従業者数} \times 100$	%
	manufacture 【従業者割合(製造業)】	$\text{従業者数(製造業)} / \text{従業者数} \times 100$	%
	telecommunication 【従業者割合(情報通信業)】	$\text{従業者数(情報通信業)} / \text{従業者数} \times 100$	%
	transport 【従業者割合(運輸業、郵便業)】	$\text{従業者数(運輸業、郵便業)} / \text{従業者数} \times 100$	%
	wholesale 【従業者割合(卸売業、小売業)】	$\text{従業者数(卸売業、小売業)} / \text{従業者数} \times 100$	%
	academy 【従業者割合(学術研究、専門・技術サービス)】	$\text{従業者数(学術研究、専門・技術サービス)} / \text{従業者数} \times 100$	%
	foodservice 【従業者割合(宿泊業、飲食サービス業)】	$\text{従業者数(宿泊業、飲食サービス業)} / \text{従業者数} \times 100$	%
	education 【従業者割合(教育、学習支援業)】	$\text{従業者数(教育、学習支援業)} / \text{従業者数} \times 100$	%
	medical 【従業者割合(医療、福祉)】	$\text{従業者数(医療、福祉)} / \text{従業者数} \times 100$	%
	service 【従業者割合(サービス業(他に分類されないもの))】	$\text{従業者数(サービス業(他に分類されないもの))} / \text{従業者数} \times 100$	%
	farmer 【農家数(販売農家(人口10万人あたり))】	$\log(\text{農家数(販売農家)} / \text{総人口} \times 100000 + 1)$	%
行政基盤	current_balance 【経常収支比率】	経常収支比率(市町村財政)	%
	civil_engineering 【歳出における土木費割合】	$\text{土木費(市町村財政)} / \text{歳出決算総額(市町村財政)} \times 100$	%
	education_expence 【歳出における教育費割合】	$\text{教育費(市町村財政)} / \text{歳出決算総額(市町村財政)} \times 100$	%
	tax_income 【課税対象所得(納税義務者一人あたり)】	$\log(\text{課税対象所得(納税義務者1人あたり)})$	千円
教育	college_graduate 【学歴割合(大卒)】	$\text{最終学歴人口(大学・大学院)} / \text{最終学歴人口(卒業者総数)} \times 100$	人
労働	unemployment 【完全失業率】	$\text{完全失業者数} / (\text{完全失業者数} + \text{就業者数}) \times 100$	%
文化・スポーツ	library 【図書館数(10万人あたり)】	$\log(\text{図書館数} / \text{総人口} \times 100000 + 1)$	%
健康・医療	doctor 【医師数(10万人あたり)】	$\log(\text{医師数} / \text{総人口} \times 100000 + 1)$	人
福祉・社会保障	nursery 【保育所数(10万人あたり)】	$\log(\text{保育所等数} / \text{総人口} \times 100000 + 1)$	%
安全	crime 【刑法犯認知件数(10万人あたり)】	$\log(\text{刑法犯認知件数} / \text{総人口} \times 100000 + 1)$	件

図 2：Best Subset Selection で変数選択を行った結果



の絶対値の大小で目的変数との関係の強さを比較できるが、人口密度の絶対値が一番大きく、従業者割合（製造業）、課税対象所得、未婚者割合、単独世帯割合、農家数、従業者割合（宿泊・飲食）、完全失業率、従業者割合（農・林業）がそれに続く。

4.2 Best Subset Selection

図2はBest Subset Selectionで全29の説明変数から C_p 基準で変数選択を行った結果を示している。図では横軸が説明変数、縦軸が C_p を表しており、選択された変数は黒く塗られている。これはAICに基づく変数選択と同様な結果となることが知られている。 C_p が小さいほどいいモデルであると解釈できるため、図の最上部に位置する23個の説明変数を持つものが最良モデルとなる。この最良のモデルに含まれるのは、歳出の土木費割合、医師数、大卒者割合、従業者割合（卸売、学研究、サービス）以外の説明変数すべてである。最良モデルの各回帰係数の推定結果は表3に与えられているが、先ほどの重回帰分析と似た結果である。

表2：重回帰分析の推定結果

変数	回帰係数の推定値	標準誤差	P-値
(切片)	0.019591	0.016202	0.226766
総人口	0.050128	0.030577	0.101313
自然人口増減率	-0.10529	0.027019	0.000101
社会的人口増減率	-0.03223	0.019922	0.105888
核家族世帯割合	-0.08347	0.024755	0.000763
単独世帯割合	0.156925	0.025995	1.93E-09
未婚者割合	0.196142	0.029965	7.82E-11
人口密度	0.375057	0.045289	2.44E-16
課税対象所得	0.236293	0.034306	7.95E-12
農家数	0.141056	0.031074	6.04E-06
経常収支比率	-0.032771	0.020134	0.103782
歳出における土木費割合	0.002082	0.016932	0.902164
歳出における教育費割合	0.047338	0.017681	0.007494
完全失業率	-0.123417	0.02211	2.76E-08
医師数	0.005705	0.021682	0.792476
学歴割合(大卒)	-0.026888	0.039635	0.497624
刑法犯認知件数	0.059249	0.022335	0.008059
図書館数	0.043784	0.016747	0.009019
保育所等数	0.025764	0.017921	0.150709
従業者割合(農業、林業)	0.120916	0.024792	1.18E-06
従業者割合(建設業)	-0.038802	0.02881	0.178217
従業者割合(製造業)	0.287207	0.048016	2.69E-09
従業者割合(情報通信業)	-0.047043	0.022529	0.036935
従業者割合(運輸業、郵便業)	0.04116	0.021301	0.053491
従業者割合(宿泊、飲食サービス)	0.010875	0.027886	0.696603
従業者割合(学術研究、専門・技術サービス)	-0.005469	0.01969	0.78124
従業者割合(宿泊業、飲食サービス業)	0.137586	0.030853	8.76E-06
従業者割合(教育、学習支援業)	-0.038387	0.022386	0.08657
従業者割合(医療、福祉)	-0.067113	0.031767	0.034773
従業者割合(サービス業(他に分類されないもの))	-0.011944	0.021216	0.573534

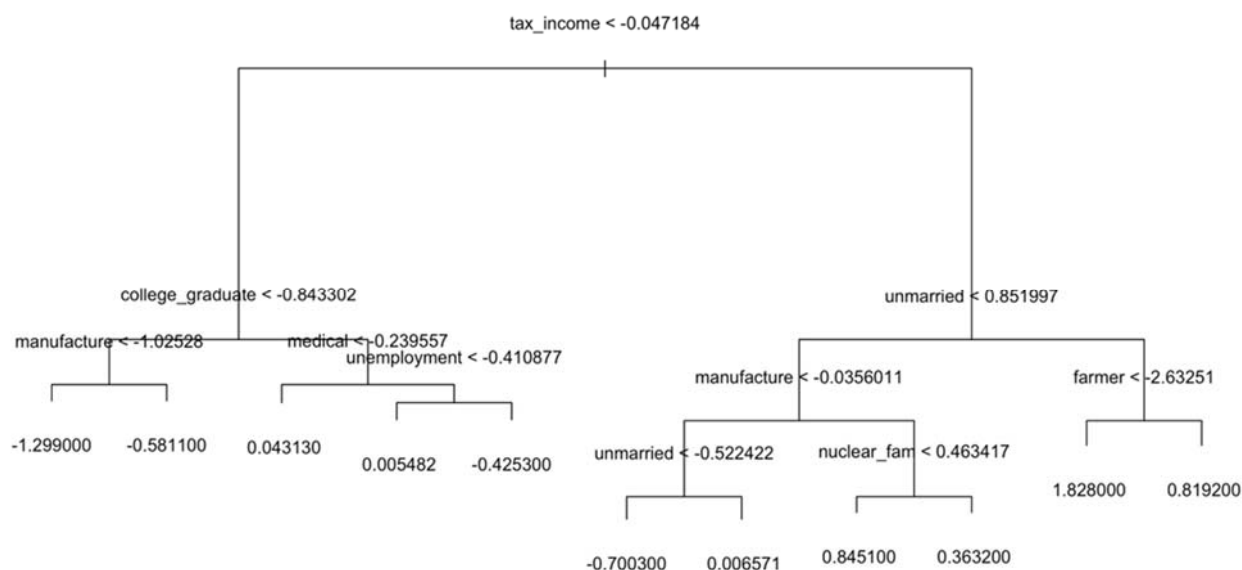
表3：Best Subset Selectionで
選択された変数の回帰係数

変数	回帰係数の推定値	標準誤差	P-値
(切片)	0.01964	0.01618	0.225061
総人口	0.05341	0.02872	0.063134
自然人口増減率	-0.10425	0.02671	9.86E-05
社会的人口増減率	-0.03274	0.01969	0.096508
核家族世帯割合	-0.08717	0.02432	0.000348
単独世帯割合	0.15281	0.02499	1.21E-09
未婚者割合	0.1923	0.02965	1.15E-10
人口密度	0.36583	0.03936	< 2e-16
課税対象割合	0.22234	0.02948	7.46E-14
農家数	0.13817	0.03045	6.10E-06
経常収支比率	-0.0332	0.02003	0.097471
歳出における土木費割合	-	-	-
歳出における教育費割合	0.04615	0.01738	0.007987
完全失業率	-0.11991	0.02152	2.94E-08
医師数	-	-	-
学歴割合(大卒)	-	-	-
刑法犯認知件数	0.05882	0.02216	0.008024
図書館数	0.04399	0.0166	0.008116
保育所等数	0.0272	0.01776	0.125855
従業者割合(農業、林業)	0.12009	0.02248	1.05E-07
従業者割合(建設業)	-0.03833	0.02447	0.11747
従業者割合(製造業)	0.28623	0.03081	< 2e-16
従業者割合(情報通信業)	-0.04972	0.02231	0.025939
従業者割合(運輸業、郵便業)	0.04156	0.01964	0.034443
従業者割合(卸売業、小売業)	-	-	-
従業者割合(学術研究、専門・技術サービス)	-	-	-
従業者割合(宿泊業、飲食サービス業)	0.13664	0.02408	1.64E-08
従業者割合(教育、学習支援業)	-0.04143	0.01951	0.033883
従業者割合(医療、福祉)	-0.06374	0.02508	0.011126
従業者割合(サービス業(他に分類されないもの))	-	-	-

4.3 決定木分析

図3には推定された決定木の構造が示されている。図の上にある変数ほど早い段階でデータ分割に利用されるので重要度が高い。最初に課税対象所得に基づいてデータが分割され、次に大卒割合、未婚割合に基づき分割が進む。他に使用された説明変数は従業者割合（製造業、医療・福祉）、農家数、完全失業率、核家族世帯割合である。推定された木構造によると、最も外国人人口が多くなるのは、所得がおよそ平均以上、未婚割合が高く、農家数がかなり低いグループで、逆に最も少なくなるのは所得が平均以下、大卒割合が低く、製造業割合も低いケースである。

図3：回帰木分析によって推定された木



4.4 ランダムフォレスト

ランダムフォレストは複数の木の平均を計算することで予測の改善を行うが、決定木の長所であるモデルの説明力の高さは失われ、データが持つ構造は明確でない。しかし、変数をモデルから外した場合にどの程度予測が悪化するかに基づいて、各々の説明変数の重要度は計算できる。図4はランダムフォレストによる回帰分析で得られた各説明変数の重要度を表している。課税対象所得の重要度が一番高く、その他に未婚割合、従業者割合（製造業、医療・福祉）、犯罪認知件数、核家族割合、完全失業率、大卒割合、人口密度などの重要度が比較的高い。重回帰分析と木に基づく手法の推定結果の大きな違いとして、大卒割合が前者では選択されない一方で、後者では重要な変数と認識されることが挙げられるが、大卒割合と外国人人口の間には非線形な関係性が存在するのかもしれない。

4.5 10-fold クロスバリデーションでの比較

最後にクロスバリデーションを行い、これまでの統計手法の妥当性を比較する。10-fold クロスバリデーションによって求められた各分析手法における平均二乗誤差の平均値はそれぞれ表4のようになった。決定木の値が最も悪いが、先ほどの結果が示すように、他の手法よりも使用する説明変数の数が少ない傾向があり、このデータにおける構造を捉えるには不十分な可能性がある。その次が線形回帰モデルであるが、Best Subset Selectionの方が説明変数の選択を行なっているため予測が少し改善している。このデータに対してはランダムフォレストが最も妥当性が高い結果となっているが、Best Subset Selectionと比較してそれほど大きな違いはないことを考慮すると、線形回帰モデルでもデータの関係性をある程度うまく捉えられてい

図 4：ランダムフォレストによる各説明変数の重要度

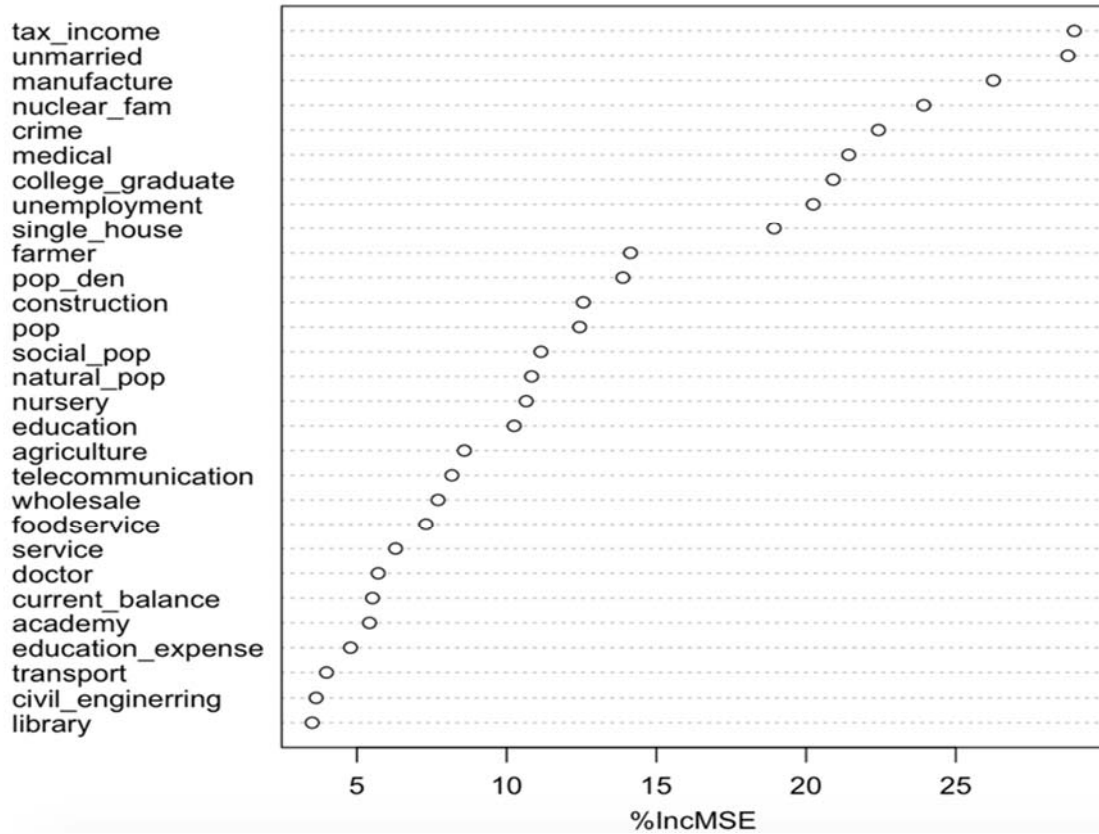


表 4：10-fold クロスバリデーションによる平均二乗誤差の平均値

重回帰分析	Best Subset Selection	決定木分析	ランダムフォレスト
0.4686977	0.4641824	0.527091	0.4305073

るのではないかと考えられる。

すべての手法で重要度が高いと考えられる変数としては、従業者割合（製造業）、課税対象所得、未婚者割合、農家数、完全失業率などが挙げられる。人口密度は決定木で選択されていないが、線形回帰モデルでは重要度が最も高いことに加え、クロスバリデーションにおける決定木の値の悪さを考慮すると、人口密度も重要度が高い説明変数だと考えられる。これらの結果から大まかに、都市化（人口密度、課税対象所得、未婚者割合）、働き手不足（完全失業率）、特定産業（農業、製造業）が外国人人口と比較的強い結びつきがあるのではないかと解釈できる。

5. 結論

本研究では外国人人口と自治体の特性の関係性について統計分析を行った。主に SSDSE-2019A の市区町村データの項目から目的変数（10 万人当たり外国人人口）と多面的な特性を表す説明変数を形成した。標準的アプローチである重回帰分析を採用し、さらに説明変数の数が 30 近くあることを考慮に入れ、データへの過適合を防ぐために Best Subset Selection により変数選択を行った。加えて、線型性を仮定しないアプロー

チとして近年広く用いられている決定木とランダムフォレストを用いて目的変数と説明変数の関係性を探った。結果として、都市化、働き手不足、特定の産業が外国人人口との結びつきが強そうだとということがわかった。これらの特性の変化を注視することで、今後の外国人人口の増減の傾向が掴める可能性があり、特に増加すると予測される市区町村にとっては外国人のための生活環境の改善に取り組むきっかけになるかもしれない。

また、本研究に関しては今後の課題も複数存在する。最初に、本研究では統計学的手法に基づいて外国人人口と関連がありそうな自治体の特性を選び出したが、この結果に対する社会的・人口学的理論の裏づけについて議論しておらず、これらの分野で長年蓄積された研究成果と整合的であるかどうかを今後検証する必要がある。また、表 1 が示すように本研究では比較的多くの説明変数を使用しているが、依然として考慮に入れるべき重要な変数を除外している可能性がある。例えば、自治体内に大学などの高等教育・研究機関があれば留学生が多いと考えられ、さらに一般に低収入の学生は生活費を抑えるために家賃水準が低い地域に集まるかもしれないが、本研究の説明変数にはこれらの情報は直接的には含まれていない。すると、大学の有無や家賃水準などの変数を含めた場合に、それらと相関がありそうな大卒割合や犯罪認知件数などの変数と外国人人口との関係性に変化が起こる可能性があり、結果の解釈を慎重に行う必要がある。

また、本研究で調べた関係性は相関関係であって因果関係ではないことにも注意が必要である。例えば、歳出における教育費割合を変えるなど、ある特定の政策が外国人人口に与える影響を分析するためには傾向スコアなどの因果推論に適切な統計手法を用いる必要がある⁽¹⁵⁾。加えて、社会構造は時間と共に少しずつ変化すると考えられる一方で、本研究ではそれぞれの変数に対して一時点のデータのみを使用しているため、将来的な構造変化について本研究から言及することには限界がある。一つのアプローチとしては、複数の異なる時点におけるデータセットを作成し、時間的な構造を明示的に取り込んだ統計手法を使って構造変化の分析を行うことが考えられる。

最後に、今回の分析に用いた外国人人口は国勢調査によるものであるが、外国人人口は在留外国人統計でも調査されている。国勢調査と在留外国人統計に記載されている外国人人数は、2015 年時点で、後者を 100% とすると前者は 78.5% となっており、両統計において外国人人数が大きく異なっている⁽¹⁶⁾ ことに注意が必要である。

参考文献

- (1) 朝日新聞デジタル「日本に住む外国人、初の 2% 超え 島根・鹿児島 15% 増」2019 年 7 月 10 日
<https://www.asahi.com/articles/ASM7B4QJ3M7BUTFK00D.html>
- (2) 日本経済新聞「人口減少幅最大の 43 万人、外国人は最多 人口動態調査」2019 年 7 月 10 日
<https://www.nikkei.com/article/DGXMZ047178730Q9A710C1EA1000/>
- (3) 日本経済新聞「外国人受け入れ拡大の制度設計を急げ」2018 年 10 月 18 日
<https://www.nikkei.com/article/DGXKZ036619070Y8A011C1EA1000/>
- (4) 石川義孝、竹下修子、花岡和聖：“2005-2010 年における新規流入移動と国内移動からみた外国人の目的地選択”、京都大学文学部研究紀要、53 巻、pp. 293-318 (2014) .
- (5) 清水昌人：“市区町村における外国人の社会増加と日本人の社会減少”、E-journal GEO、12 巻、1 号、pp. 85-100 (2017) .
- (6) 山内昌和：“近年の日本における外国人女性の出生数と出生率”、人口問題研究、66 巻、4 号、pp. 41-59 (2010) .
- (7) 是川夕：“日本における外国人女性の出生力 一国勢調査個票データによる分析”、人口問題研究、69

卷、4号、pp. 86-102 (2013) .

- (8) 是川夕：“外国人の居住地選択におけるエスニック・ネットワークの役割”、社会学評論、59 卷、3 号、pp. 495-513 (2008) .
- (9) 中村二郎：“外国人労働者の受け入れは何をもたらすのか”、日本労働研究雑誌、587 号、pp. 16-26 (2009) .
- (10) 中川雅貴、山内昌和、菅桂太、鎌田健司、小池司朗：“都道府県別にみた外国人の自然動態”、人口問題研究、74 卷、4 号、pp. 293-319 (2018) .
- (11) James G.、Witten D.、Hastie T.、Tibshirani R.：“An Introduction to Statistical Learning: with Applications in R”、Springer (2017) .
- (12) 「都道府県・市区町村のすがた(社会・人口統計体系)」e-Stat 政府統計の総合窓口 <https://www.e-stat.go.jp/regional-statistics/ssdsview>
- (13) 前田泰伸：“都道府県別に見た外国人労働者と経済の関係 ～経済的に好調な都道府県に外国人労働者も集まる～”、経済のプリズム、177 号、pp. 11-19 (2019) .
- (14) 社会・人口統計体系の中の項目定義のサイト
https://www.e-stat.go.jp/koumoku/koumoku_teigi/C#C3102
- (15) 星野嵩宏：“調査観察データの統計科学 因果推論・選択バイアス・データ融合”、岩波書店 (2009) .
- (16) 石川義孝：“地図で見る日本の外国人 改訂版”、p. 2-3、ナカニシヤ出版 (2019) .