

離島データの識別情報に基づく分析

201011368 吉元 翔汰

経営工学主専攻 指導教員：イリチュ（佐藤）美佳 教授

1. はじめに

現代の日本では、都会の過密、地方の過疎化、それに伴う都会と地方の格差が問題となっている。特に地方では、人口減少、少子高齢化、雇用の減少など多くの問題が起っている。日本の経済を支えてきた地方の産業は、過疎化とともに衰退し「限界集落」と呼ばれる人口の50%以上が65歳以上の高齢者で社会的共同生活の維持が困難になっている集落まで出てきている。そのような中で、日本の水産業を支えてきた多くの離島は過疎化の影響で、一つの自治体として維持していくことが難しくなっている。また、日本の離島のほぼすべてが過疎化の悩みを持ち、活性化の方法を模索している。そこで、日本の中でも特に過疎化が進んでいる離島の人口減少の実態を解析することによって、過疎化を食い止めるためにはどのような政策が必要となるのかがわかるのではないかと考えた。最終的には、離島の過疎化の要因がわかることによって、離島だけでなくこれから日本のあらゆる所で起こるであろう過度の過疎化に対する対応策を考える上でも役立つのではないかと考えた。

離島のデータに関する研究はあまり行われていなく先行研究は少ない。そのような中で、規模・基盤・産業・行政政策の経年変化にみる離島構造特性と類型化を行っている先行研究 [2] もあるが、1999年から2010年まで政府主導で行われた市町村合併である平成の大合併の影響等を含めた離島の人口に対する詳細な分析はされていない。そこで、本研究では2011年の『離島統計年報』[3]のデータや国勢調査 [4][5][6]等のデータから、現在日本全体に307島ある有人離島の人口のデータ、地理的なデータや産業のデータなどを抽出し、分類構造を考慮に入れたブロッキング回帰分析 [1] を行い、離島のデータに対するブロッキング回帰分析の有効性の検討と離島の人口減少の実態を解析することを目的とした。

2. 離島の現状

図1から離島の人口が年々減少し、2000年から2010年の間に12.5%減少していることがわかる。全国の人口が0.9%増加していることと比べると急速に離島の人口が減少していることがわかる。また、図2から全国の高齢化率が23.1%である中で、離島の高齢化率は33.0%とより高齢化が進行している。図3から、20歳から39歳までの人口の割合は全国と比べて異常に低い値を示していることがわかり、このことから若年層の人口の流出が進んでいることが考えられる。このように、離島では過度に過疎化が進んで

いることがわかる。

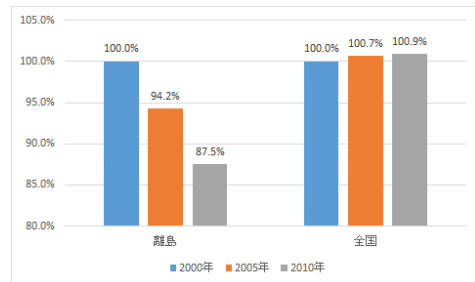


図1 離島と全国の人口動態の比較 (2000年の人口が基準)

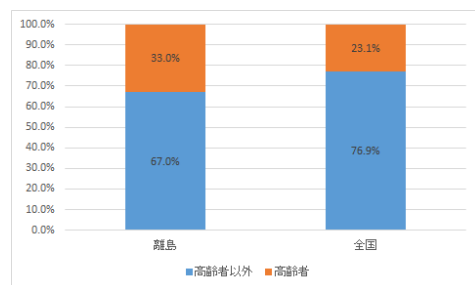


図2 離島と全国の高齢者比率の比較

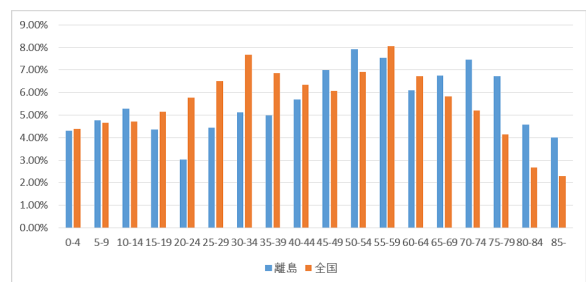


図3 離島と全国の年齢別人口の割合の比較

3. ブロッキング回帰分析

ブロッキング回帰分析 [1] とは、外部の分類情報に基づきデータを各ブロックに分類し、そのブロックごとの違いを表すダミー変数を含めて回帰モデルを作成する分析方法である。

個体が n 個、変数が k 個のデータ行列を X 、 r 個のブロックを表す $n \times (r-1)$ の行列を \tilde{X} としたとき、 $n \times (k+r)$ の

行列 \hat{X} を

$$\hat{X} = [X_0 | \tilde{X} | X] \quad (1)$$

とする。ここで、 X_0 はすべての要素が1のベクトルである。そして、 Y を目的変数とするブロック回帰モデルは、行列 \hat{X} と回帰係数 β と残差 ε を用いて、

$$Y = \hat{X}\beta + \varepsilon \quad (2)$$

と定義される。

r 個のブロックに分けられるとき、ブロックを表すダミー変数は $r-1$ 個必要となる。例えば、説明変数が2つ、ブロックが3つのブロック回帰モデルは、

$$Y = \beta_0 \hat{x}_0 + \beta_1 \hat{x}_1 + \beta_2 \hat{x}_2 + \beta_3 \hat{x}_3 + \beta_4 \hat{x}_4 + \varepsilon \\ = \beta_0 + \beta_1 \tilde{x}_1 + \beta_2 \tilde{x}_2 + \beta_3 x_1 + \beta_4 x_2 + \varepsilon \quad (3)$$

と表され、 \tilde{x}_1, \tilde{x}_2 がブロックを表す行列、 x_1, x_2 は説明変数である。式(3)において $\beta_0, \beta_1, \beta_2$ の値を比べることで、残りの説明変数の影響をなくしたブロックごとの違いを見ることができる。

4. 数値例

離島振興法、小笠原諸島振興開発特別措置法、奄美群島振興開発特別措置法、沖縄振興特別措置法の各法に基づき指定されている日本の離島307島のデータについて解析を行った。このデータの出典は、日本離島センターが毎年発行している『離島統計年報』[3]に掲載されているデータと国勢調査のデータである。以下がデータの内容である。

データの内容

概要、人口・世帯数・人口動態、年齢別男女別人口、産業分類別就業者数、保育所・学校現況、農林業現況、水産業現況、農林水産業生産額、観光客数、宿泊能力、自然公園、道路現況、車輛保有台数、港湾・航路現況、空港・航空路現況、医療施設・医療従事者等現況、一般廃棄物処理状況(し尿処理)、一般廃棄物処理状況(ごみ処理)、汚水処理状況

まず、解析にあたり以下の離島のデータを除いた。

- 平成25年4月施行の離島振興法改正につき離島指定を解除された離島
- 総人口が100人以下の離島(規模が小さい)
- 宮城県の離島(欠損値が多い)

次に、回帰モデルによって離島の人口減少の要因を分析するために、次のように人口に関わる変数を目的変数の候補とし、人口に影響のありそうな変数を説明変数の候補とした。

- 目的変数
総人口、生産年齢人口(15歳から64歳までの人口)、20代と30代の人口
- 説明変数
総面積、農業生産額、水産業生産額、観光客数、自然公園の面積、医師数、病院数、保有車両総数、第1次産業

就業者数、第2次産業就業者数、第3次産業就業者数、医療福祉関係就業者数、教育関係就業者数、公務員数

まず、総人口、生産年齢人口、20代と30代の人口というそれぞれの目的変数に対して重回帰分析を行った。ここでの変数選択には、AICの値を基準にした変数増減によるステップワイズ法を用いた。次に、重回帰モデルに市町村合併の有無や高校の有無というダミー変数を加えてブロック回帰分析を行った。その結果が、表1である。目的変数ごとに、決定係数(R^2)、自由度修正済決定係数(*Adjusted* R^2)、AICの値すべてが良い値をとっているものを見ていくと、総人口ではモデル3-1、生産年齢人口ではモデル3-2、20代と30代の人口では2-3であり、すべてブロック回帰モデルである。

表1 モデルごとの決定係数、自由度修正済決定係数、AICの値の比較

モデル	目的変数	ブロック	R^2	<i>Adjusted</i> R^2	AIC
1-1	総人口	なし	0.9967	0.9966	2189
2-1		合併の有無	0.9967	0.9966	2191
3-1		高校の有無	0.9971	0.9970	2172
1-2	生産年齢人口	なし	0.9947	0.9946	2086
2-2		合併の有無	0.9948	0.9946	2086
3-2		高校の有無	0.9952	0.9951	2073
1-3	20代と30代の人口	なし	0.9782	0.9777	1978
2-3		合併の有無	0.9788	0.9782	1976
3-3		高校の有無	0.9782	0.9776	1980

5. まとめ

重回帰分析とブロック回帰分析の結果比較より、ブロック回帰分析の方がよい結果を得た。このことから、離島のデータに対して市町村合併の有無や高校の有無を考慮したブロック回帰分析が有効であるということがわかった。また、このことから市町村合併の有無や高校の有無ということが、離島の人口に対して何らかの影響を与えているのではないかと考えられる。

参考文献

- [1] N.R. Draper and H. Smith, Applied Regression Analysis, John Wiley and Sons, pp.134-142, 1966.
- [2] 姫野由香, 牧田正裕, 規模・基盤・産業・行政政策の経年変化にみる離島構造特性と類型化, 平成21年度国土政策関係研究支援事業 研究成果報告書, 2009.
- [3] 2011 離島統計年報 CD-ROM 版, 財団法人 日本離島センター, 2013.
- [4] 平成12年国勢調査, 総務省統計局, 2000.
- [5] 平成17年国勢調査, 総務省統計局, 2005.
- [6] 平成22年国勢調査, 総務省統計局, 2010.