

# 都道府県別人口動態に関する統計分析

2020SS033 前田憲希

指導教員：白石高章

## 1 はじめに

現在日本では少子高齢化が深刻な問題とされている。そのため、教育現場や医療現場など様々な現場で人手不足問題が起きている。私たちが大学を卒業し社会人になった時、若者一人につき約2人の高齢者を支えなければならないことを知った。このままでは、子供の数が減り、日本が老人だらけになってしまうという不安感や、私たち若者がこの先安心して生活を送れるのかを知りたいと思い、人口動態についての統計を行いたいと思った。

本研究では出生率と転入超過数を調べることで人口動態を調べていく。この2つの要素を選んだ理由として、出生率が高いと子供の数が多く、低いと子供の数が少ないという解釈を都道府県別に解釈することができる。子供の数が増えると日本の人口が増える。そのため、人口の増減で人口動態を調べることができると思い、出生率を選択した。

また、転入超過数が正の数であれば、その都道府県は人口が増える。逆に負の数であれば、その都道府県は人口が減る。そのため、人口の移動という人口動態を調べることができると思い、転入超過数を選択した。

## 2 データ

データは令和元年の estat の政府統計の総合窓口から引用したデータと都道府県が提示しているデータを引用した。また、これらのデータは令和元年に統一した。 $x_1$ =(人口),  $x_2$ =(出生率),  $x_3$ =(死亡率),  $x_4$ =(婚姻率),  $x_5$ =(離婚率),  $x_6$ =(面積),  $x_7$ =(人口密度),  $x_8$ =(犯罪率),  $x_9$ =(犯罪率),  $x_{10}$ =(平均気温),  $x_{11}$ =(平均湿度),  $x_{12}$ =(年間降水量),  $x_{13}$ =(1年間のお酒の消費量),  $x_{14}$ =(転入超過数),  $x_{15}$ =(最低賃金),  $x_{16}$ =(一年間の交通事故件数),  $x_{17}$ =(大学の数),  $x_{18}$ =(喫煙率) とする。

分析方法は相関係数での分析、クラスター分析、重回帰分析で行った。これらの分析は主に「1」を参考に行った。

## 3 相関係数

出生数、転入超過数と様々なデータをとの相関係数を調べた。相関係数については本研究では±0.4から相関があると定義し研究を行った。その後、相関がある変数に対して有意水準を0.05、帰無仮説を「2つの変数間に相関がない」とし、無相関検定を行った。相関関係があると考えた2変数のうち帰無仮説を棄却するものを相関関係があるとする。相関係数表を可視化した表を図1に示す。図1に示されているAからQの並びはデータについての $x_1$ から $x_{17}$ までの並びと同じである。この表は正の相関であれば青く、負の相関なら赤くなる。また、相関が強ければ強いほど色が濃くなる。

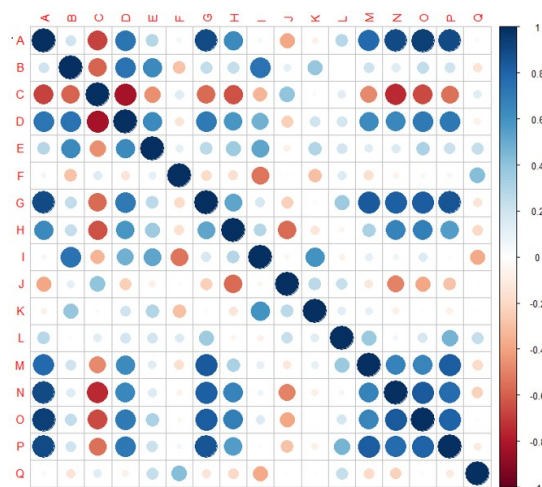


図1 相関係数表 可視化

図1から出生率と正の相関があった要素は婚姻率(0.74), 離婚率(0.64), 平均気温(0.73)。負の相関は死亡率(-0.59)であった。また、転入超過数と正の相関があった要素は人口(0.78), 婚姻率(0.62), 人口密度(0.84), 最低賃金(0.67), 一年間の交通事故数(0.65) 大学の数(0.82) 死亡率(-0.47)であった。また、これらの結果はp値が有機水準0.05をすべての要素が下回っていたため、相関がある。出生率は婚姻率と強い正の相関があった。出生率と離婚率, 平均気温が正の相関があるのは意外であった。また、転入超過数では正の相関がある要素から都会であるほど人口が集中している。死亡率は、出生率、転入超過数と共通して負の相関であった。

## 4 クラスター分析

18個のデータを標準化してクラスター分析をウォード法を用いて行った。また、東京都の人口が与える影響が大きいため、人口のデータを消し、クラスター分析を行った。

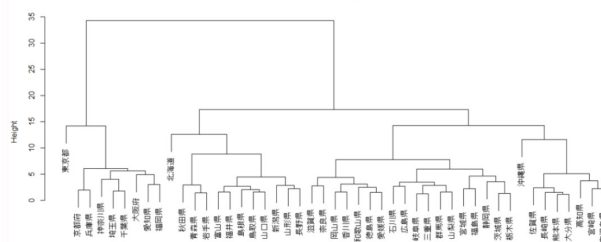


図2 解析結果

図2より,47都道府県を4つの群に分けた。  
 第一群：人口密度が他の地域よりも低い地域  
 この群の特徴は、面積は大きいが人口が少ないため、人口密度が非常に小さい。また、この群の中に人口が少ないと言われている鳥取県、島根県がある。

第二群：都市部郊外にあり、生活がしやすい地域  
 この群に含まれている地域はすべて郊外の県である。特徴として、人口密度がそんなに高くなく、街も一部栄えているためとても生活しやすい地域であることが分かる。

第三群：都会と言われる地域  
 この群では、すべてが都市部で、人口が集中している地域である。面積のわりに全国から人が集まっているため、常に人口密度が高い状態を保っている。

第四群：田舎であるが、子供が多い地域  
 この群は人口が少なく、田舎といわれる地域が多い。しかし、データを見ると、この群に含まれている地域は、出生数と死亡数の差があまりない。このことから、人口は少ないが子供を産みやすい地域であることが分かった。

## 5 重回帰分析

次に、出生率と転入超過数が何と関係しているのかを調べるため、重回帰分析を行った。減少法で変数選択を行った。減少法によると変数は、死亡率、婚姻率、平均気温、平均湿度、最低賃金の5つとなった。変数選択後の重回帰分析の結果は次の表1である。

表1 出生率を目的関数

説明変数	回帰係数	標準誤差	p 値
死亡率	-0.4538	0.1050	0.0000
婚姻率	0.5540	0.1027	0.0000
平均気温	0.3629	0.0627	0.0000
平均湿度	0.1852	0.0608	0.0040
最低賃金	-0.4956	0.0880	0.0000

分析の結果,5つのデータのp値が0.05を下回り、標準誤差も0に近い数値であった。しかし平均気温、平均湿度は沖縄の影響が強く出ていると思った。そのため、沖縄県のデータを除いて46個のデータで重回帰分析を行った。その結果が表2である。

表2 沖縄のデータを取り除いた重回帰分析の結果

説明変数	回帰係数	標準誤差	p 値
死亡率	-0.5259	0.1447	0.0008
婚姻率	0.6661	0.1307	0.0000
平均気温	0.4018	0.0764	0.0000
平均湿度	0.2194	0.0825	0.0112
最低賃金	-0.5923	0.1325	0.0000

この結果から47都道府県を用いた重回帰分析の結果と同じ要素がでた。そのため平均気温、平均湿度は沖縄県に

影響されず、関係があることが分かった。また、表3は転入超過数を目的変数として重回帰分析を行った結果である。

表3 転入超過数を目的関数

説明変数	回帰係数	標準誤差	p 値
人口密度	0.1981	0.0365	0.0000
犯罪率	-0.1566	0.0038	0.0002
大学の数	0.1075	0.0029	0.0005

分析の結果、人口密度、犯罪率、大学の数の3つデータのp値が0.05を下回り、標準誤差も0に近い数値であった。無相関検定を行った際に死亡率は負の相関、婚姻率、人口密度、大学の数は正の相関があると分かっているため、関係性は十分にある。出生率は平均温度、平均湿度に関係していることは意外であった。データを見ると寒い地域と温かい地域では出生率に大きく差があることが分かる。このことから、出生率は平均温度、平均湿度に正の相関があると言える。最低賃金が増えれば生活に余裕ができ、出生率も上がると思う。今後の出生率を上げるために、最低賃金を底上げすることが重要であると思う。また、転入超過数は人口密度に関係していることから、犯罪率が高い地域に多くの人口が移動をするため、相対的に犯罪率は関係していると考えられる。

## 6 終わりに

出生率と転入超過数が一体何と関係しているのか私が予想していたものとあまり大差ないと思っていた。しかし、この三つの分析の結果からそんなことはなかった。すべての要素が人口に関するだけでなく社会的な要素など今から変えられる要素が関係していることが分かった。このことから、出生率を上げることは難しいことではあるが、不可能ではないことが理解できた。少子高齢化による世の中に及ぼす影響はとても大きい。例えば、労働力の減少、今後の経済や市場規模の縮小、現役世代の負担の増加、行政サービスの水準の低下など、社会経済に多大な影響を及ぼすことが分かっている。この問題を解決することは容易なことではない。そのため、行政が現在よりも子供を育てやすい環境を提供することが、この問題を解決することにつながると思われる。

## 参考文献

- [1] 金明哲:「Rによるデータサイエンス データ解析の基礎から最新手法まで」
- [2] 都道府県市区町村 <https://uub.jp/> 閲覧日 2024年1月12日
- [3] e-Stat 政府統計の総合窓口 <https://www.e-stat.go.jp/> 閲覧日 2023年1月12日
- [4] 東京新聞 <https://www.tokyo-np.co.jp/article/233787> 閲覧日 2024年1月12日