

2020年度 統計データ分析コンペティション  
特別賞（統計分析） [高校生の部]

空き家増加問題を相関関係から読み解く

好田駿成（学習院高等科）

論文の概要

社会問題にもなっている空き家増加問題に関して、空き家率と関係があると思われる変数(高齢化等に関するもの、人口密度等に関するもの、実質公債費率)を使用して空き家率との相関関係の仮説検定を行い、結果として、空き家率との相関性が高いのは高齢化率であることを見出している。

論文審査会コメント

SSDSEを利用して実証型の論文に要求される構成に準拠している。

# 空き家増加問題を相関関係から読み解く

好田駿成\*1

\*1学習院高等科

## 要旨

近年、社会問題にもなっている空き家増加問題に関して、教育用標準データの市区町村別データから様々なデータを用い、また、平成三十年度住宅・土地統計調査データからは、空き家に関するデータを抽出し、それらの抽出データから、空き家率や高齢化率等を算出し、データの加工をおこなった。そして、算出した空き家率と関係があると思われる変数(高齢化等に関するもの、人口密度等に関するもの、実質公債費率)を使用し、統計解析プログラミング言語であるRを用いて、それら変数と空き家率との相関関係を明らかにするとともに、それぞれの相関関係について、仮説検定を実施した。その結果、空き家率との相関性が高かったのは、高齢化率であった。加えて、これらの結果を総合的に考察を加え、地方公共団体の財政健全度を向上させるなどにより、空き家増加抑止が可能であるという政策提言も行った。

## 1. 研究のテーマと目的

空き家の増加により、防災、防犯、衛生、景観など、様々な社会的問題が、近年発生しており、平成二十六年には、国土交通省において、「空家等対策の推進に関する特別措置法(空家法)」も制定され、政府による空き家対策が徐々に進行しているところである。しかしながら、平成三十年度住宅・土地統計調査(以下、住宅・土地統計調査とする)によると、全国の空き家は約849万戸にも及んでおり、それが全住宅に占める割合はおよそ13.60%となっている。これを5年前の平成二十五年のデータと比較するとおよそ30万戸増加しており、20年前の平成十年のデータと比較すると、空き家の総数はこの20年で272万戸増加しており、約1.5倍にも増加しているのが現状である。

このような増加の一途をたどっている空き家については、学術レベルで様々な研究がなされてはいるものの、各種統計データを使用して発生にかかわる要素を分析した研究論文は見当たらなかった。そこで本研究では、統計データを使って、空き家増加問題への要因分析をすることを目的とする。

## 2. 研究の方法と手順

この研究では、空き家と関係があると推測される要素、例えば、総人口に対する高齢者率、全世帯に対する核家族世帯率や可住地域面積割合などの様々な要素を、教育用標準データの市区町村別データとe-Stat(<https://www.e-stat.go.jp/>)の住宅・土地統計調査データから抽出し、Rを用いて相関関係を明らかにした。加えて、それらの相関関係に関して、以下のような仮説検定を行なった。

$\rho$  を空き家率と  $i$  の相関係数として

- 帰無仮説  $H_0: \rho_i=0$
- 対立仮説  $H_A: \rho_i>0$

または

- 帰無仮説  $H_0: \rho_i=0$
- 対立仮説  $H_A: \rho_i<0$

検定統計量は  $t = \rho_i \sqrt{(n-2) / \sqrt{1-\rho_i^2}}$ 、 $\alpha = 0.05$  とする。

### 3. データセットの加工

表1 利用した教育用標準データの市区町村別データの項目

使用変数	出典	年度
総人口(人)	総務省統計局「国勢調査報告」人口等基本集計	2015
65歳以上人口(人)	総務省統計局「国勢調査報告」人口等基本集計	2015
死亡数(人)	厚生労働省政策統括官付参事官付「人口動態統計」	2018
転入者数(人)	総務省自治財政局「地方財政統計年報」「市町村別決算状況調」	2017
転出者数(人)	総務省自治財政局「地方財政統計年報」「市町村別決算状況調」	2017
高齢単独世帯数(世帯)	総務省統計局「国勢調査報告」人口等基本集計	2015
可住地面積(ha)	国土交通省国土地理院「全国都道府県市区町村別面積調」	2018
世帯数(世帯)	総務省統計局「国勢調査報告」人口等基本集計	2015
総面積(北方地域及び竹島を除く)(ha)	国土交通省国土地理院「全国都道府県市区町村別面積調」	2018
実質公債費率(%)	総務省自治財政局「地方財政統計年報」「市町村別決算状況調」	2017

表2 利用した平成三十年度住宅・土地統計調査の項目

使用変数	出典	年度
22_空き家-0_総数(空き家総数)(戸)	平成三十年度住宅・土地統計調査	2018
0_総数-0_総数(住宅総数)(戸)	平成三十年度住宅・土地統計調査	2018

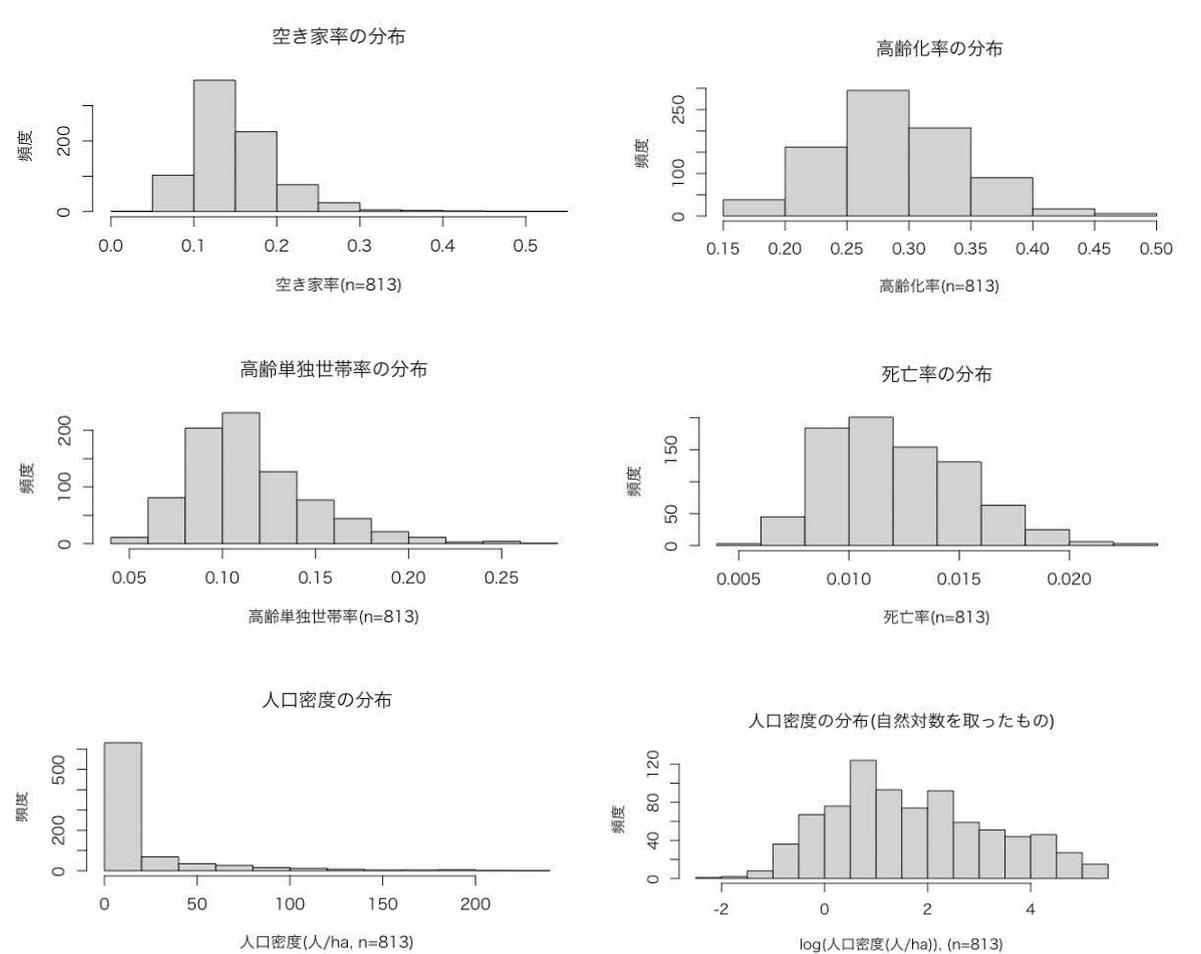
表3 利用した変数

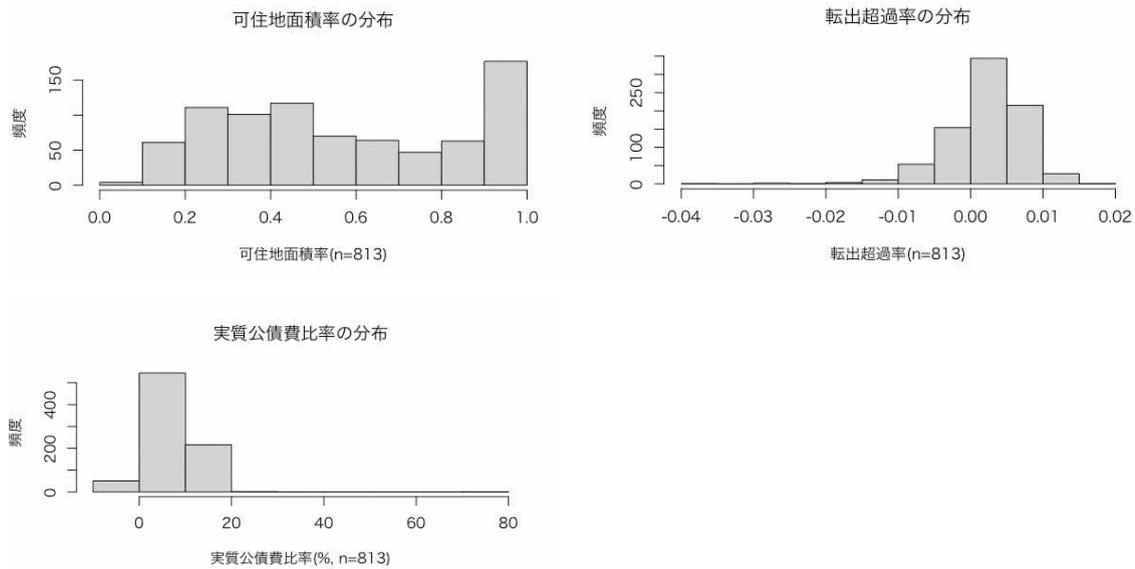
変数	定義
空き家率	空き家総数/住宅総数
高齢者率	65歳以上人口/総人口
高齢単独世帯率	高齢単独世帯数/世帯数
死亡率	死亡者/総人口

log(人口密度(人/ha))	log(総人口/総面積(北方地域及び竹島を除く)) (下のヒストグラムを見ると人口密度が偏っていたため、自然対数を用いて、偏りを解消した。)
可住地面積率	可住地面積/総面積(北方地域及び竹島を除く)
転出超過率	(転出者数-転入者数)/総人口
実質公債費率(%)	そのまま使用。

表3は、教育用標準データの市区町村別データと住宅・土地統計調査のデータは、双方の地域コードを用いて統合させた。統合させたことによって、教育用標準データの市区町村別データは、1741項目(統合前)が、813項目に減少した。

なお、使用した変数の分布は、以下のヒストグラムにまとめた。





特に、人口密度に関しては、分布が偏っていたため、自然対数を用いて偏りを解消した。

#### 4. データ分析の結果

##### ①空き家率と高齢化率との関係について

高齢者居住の住宅は、その居住者である高齢者が死亡すると、それまで居住していた住宅を利用されない場合には、空き家として存在してしまう。これが、空き家が発生するプロセスの一つであると考える。

そこで、空き家率と高齢化率には正の相関があるのではないかとの仮説を立てた（対立仮説）。そして、その相関関係を調べたところ、相関係数  $\rho_{\text{高齢者率}}$  は0.695（その際の  $t$  値は27.582、 $p$  値は  $2.2 \times 10^{-16}$ ）という結果となった（図1）。ここで、 $p$  値が  $\alpha (=0.05)$  より低いので、帰無仮説を棄却でき、 $\alpha$  水準において統計学的に有意であると言える。この結果は、上に述べたシナリオを支持するものである。

しかしながら、上記プロセスは、空き家発生プロセスをかなり単純化しているため、更に分析を深度化した。すなわち、複数の高齢者居住住宅や多世代家族による住宅においては、高齢者の死亡により、即、空き家にはならないので、高齢化率の指標だけでは、不十分であることに気づいた。したがって、高齢単身世帯（65歳以上の者一人の世帯）の割合と死亡率にさらに焦点を当て、空き家率との関係についても調べた。

その結果を以下に示す。

相関係数  $\rho_{\text{高齢単身世帯率}}$  は0.718（その際の  $t$  値は29.41、 $p$  値は  $2 \times 10^{-16}$ ）（図2）

相関係数  $\rho_{\text{死亡率}}$  は0.692（その際の  $t$  値は27.354、 $p$  値は  $2 \times 10^{-16}$ ）（図3）。

$p$  値はどちらも  $\alpha$  より低いので、帰無仮説を棄却でき、 $\alpha$  水準において統計学的に有意である。つまり、高齢単身世帯率や死亡率が増加（減少）すると空き家率も増加（減少）するということである。

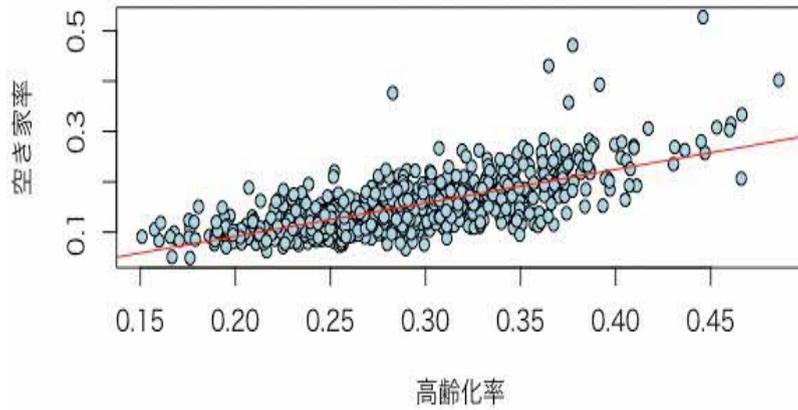


図1 高齢化率と空き家率の相関関係  
(赤線は回帰線)

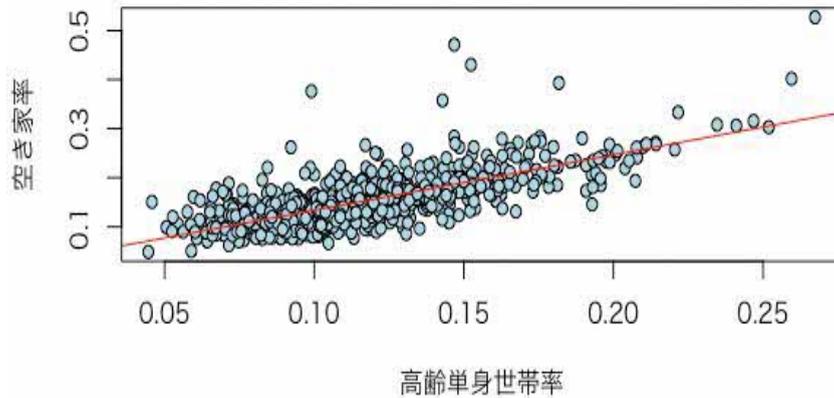


図2 高齢単身世帯率と空き家率の相関関係  
(赤線は回帰線)

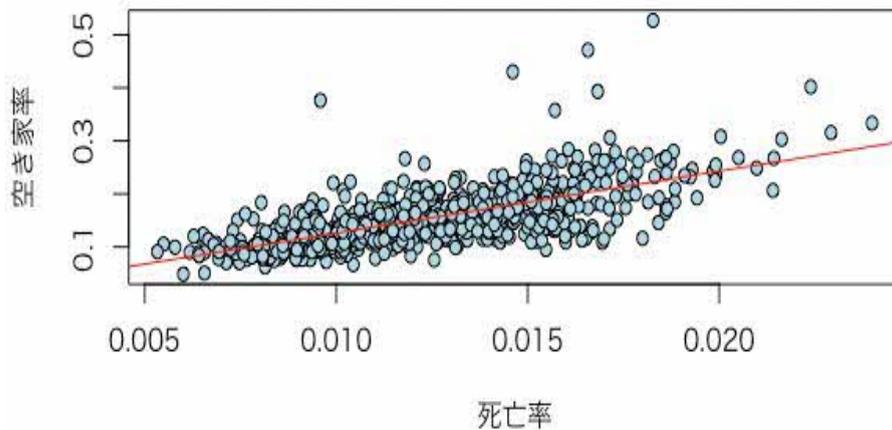


図3 死亡率と空き家率の相関関係  
(赤線は回帰線)

## ②空き家率と人口密度との関係

人口は都道府県で大きく異なり、首都である東京都では約1400万人であり、その一方、人口の少ない鳥取県の人口はおよそ56万人しかいない。また、人口密度についても、都会では高く田舎では低いという関係性がある。そこで、人口密度が低い地域(いわゆる田舎)と空き家率との間には、正の相関関係がある(対立仮説)と考えた。そこで、①と同様に、統計分析を行ったところ、相関係数 $\rho_{\text{人口密}}$

度は-0.535その際の  $t$  値は-18.08、 $p$  値は $2 \times 10^{-16}$ であった(図4)。ここに、 $p$ 値が $\alpha$ より低いので、帰無仮説を棄却でき、 $\alpha$ 水準において統計学的に有意である。この結果より、人口密度が高い(低い)と、空き家率は低く(高く)なることがわかる。

加えて、人口密度のみならず可住地面積率(総面積に対する総面積から山林や湖沼面積を差引いた可住地面積の割合)が低いほど、いわゆる田舎を表す指標であることから、可住地面積率と空き家立には、負の相関関係がある(対立仮説)と考えた。そして、同様に、相関関係を調べてみると、相関係数 $\rho_{\text{可住地面積率}}$ は-0.514(その際の  $t$  値は-17.08、 $p$  値は $2 \times 10^{-16}$ )であった(図5)。 $p$ 値が $\alpha$ より低いので、帰無仮説を棄却でき、 $\alpha$ 水準において統計学的に有意である。

上記2つの相関関係より、都会であると空き家が少なくなり、田舎であると空き家が多いという結果になった。

さらに、ほかの要素においても、上記結論が導けるかどうかとの疑問により、転入者数と転出者数の統計データより、転出超過率を算定し、それと空き家率との相関関係を調べた。その結果、相関係数 $\rho_{\text{転出超過率}}$ は0.462(その際の  $t$  値は14.85、 $p$  値は $2 \times 10^{-16}$ )であった(図6)。 $p$ 値が $\alpha$ より低いので、帰無仮説を棄却でき、 $\alpha$ 水準において統計学的に有意である。よって、転入者数より転出者数が多い(少ない)地域ほど、空き家が多い(少ない)という結果になった。

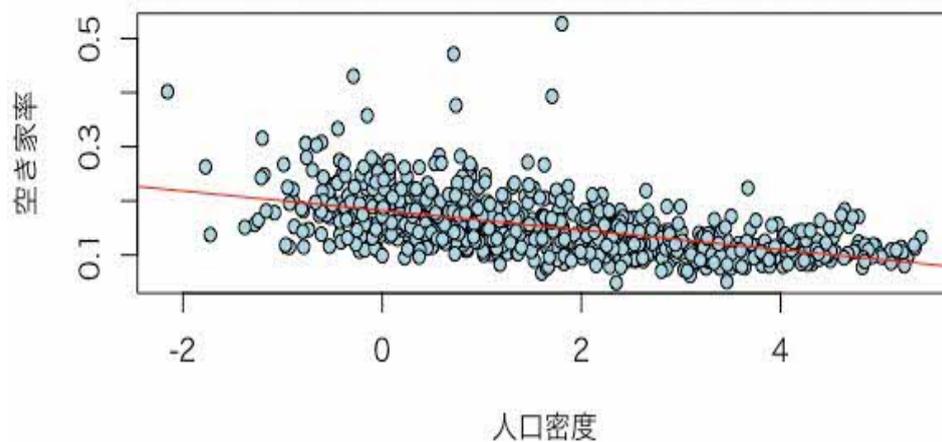


図4 人口密度と空き家率の相関関係  
(赤線は回帰線)

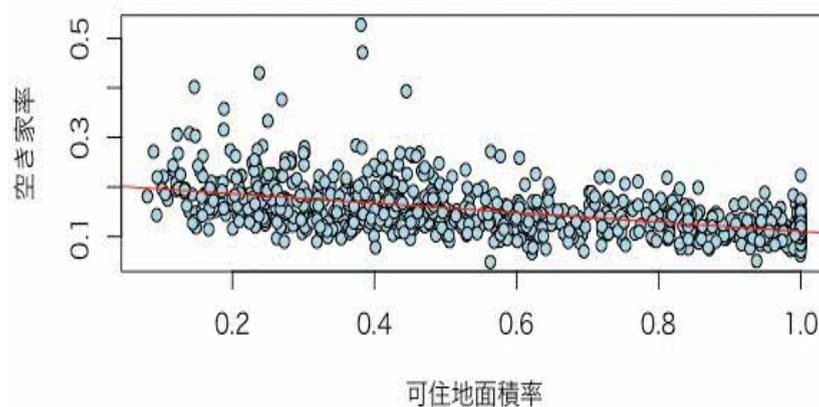


図5 可住地面積率と空き家率の相関関係  
(赤線は回帰線)

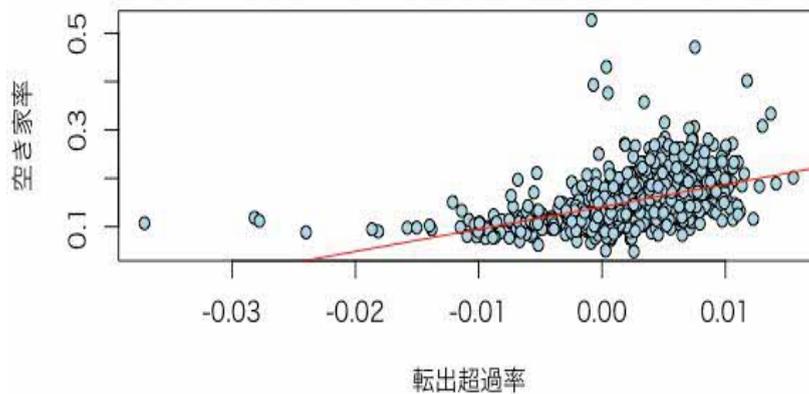


図6 転出超過率と空き家率の相関関係  
(赤線は回帰線)

### ③空き家率と実質公債費率との関係

ここでは、経済的な指標に関する相関関係の分析についても試みたので、それに関して記述する。人々が居住地を選択する場合、良好な居住環境は、判断材料としてかなりの比重を占めると考えられる。例えば、子育てがしやすい公園や学校、教育施設が充実している、道路が広い、バスなどの交通移動手段が発達している、などということが良好な居住環境であり、それを提供出来るかどうかは、自治体がどれだけ裕福であるかによると考えた。

そこで、財政健全度の指標の一つである、実質公債費率について注目し、空き家との負の相関関係があると仮説を立てた。実際に相関関係を調べたところ、相関係数 $\rho_{\text{実質公債費率}}$ は0.374(その際の t 値は11.48、p 値は $2 \times 10^{-16}$ )であった(図7)。p 値が $\alpha$ より低いので、帰無仮説を棄却でき、 $\alpha$ 水準において統計学的に有意であると言えることがわかった。このことから、実質公債費率が大きく(少く)なると空き家率大きく(少く)なることが分かった。

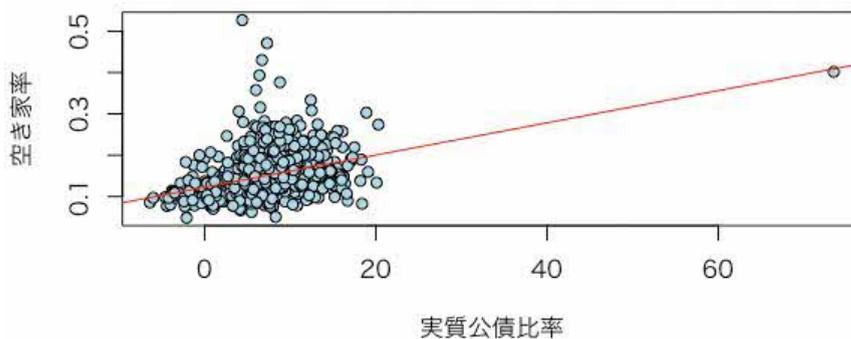


図7 実質公債比率と空き家率の相関関係  
(赤線は回帰線)

## 5. 結果の解釈

4の結果より、相関係数の観点から見ると、三つの要素のうち、最も相関性が高いのは、①の空き家率と高齢化率との関係であり、高齢化率が空き家増加の要因だと言える。加えて、①で考察した、高齢者居住の住宅は、その居住者である高齢者が死亡すると、その住宅を利用されない場合には、空き家として存在することになるというプロセスは空き家増加に関連していると思われる。また②は①と比べて、相関係数はやや弱いですが、それぞれの要素が空き家の増加に関係していると思われる。②の結果から、田舎は空き家率が多く、都心では空き家率が低いという解釈ができる。③については、②と同様に相関関係はやや弱いですが、相関関係がないとは言えない。これらの結果から、空き家率に一番関わっているのは高齢化率であるとわかった。

また、これらの結果から、以下のような政策提言が可能になるとと思われる。

まず、地方公共団体が財政健全度を向上させて、その地域の住環境を向上させる。③でも述べ、繰り返してはあがるが、子育てがしやすい広くて快適な公園や、学校、教育施設が充実していること、道路が広くて安全であること、バスなどの公共交通機関が充実していれば、魅力ある町となり、その地域の魅力が増し、転入者が増える。するとその地域が活性化し、その地域の高齢者数は変わらないがその地域の高齢化を抑制できる。そうすることによって高齢化率が減少し、この研究でわかった高齢化率と空き家率の関係から、空き家率が減少し、空き家増加の抑制ができる。

最後に、この研究の限界としては、住宅・土地統計調査に関するデータの最新版が2018年、教育用標準データの市区町村別データの変数に関しては、一番古いものでは2015年であったため、比較した2つのデータには最大3年の差があった。しかし、このデータより使った変数は3年間で大幅に変化するとは考えられず相関関係にはそこまで影響がないだろうと考えた。また、相関関係や仮説検定だけでなく、もう少し進んだ統計手法を利用した、研究もできるだろうと考える。

## 6. 参考文献

(1) 国土交通省：“令和元年5年空き家対策について”

<https://www.mlit.go.jp/common/001290020.pdf>

(2) 総務省：“平成三十年度住宅・土地統計調査”

<https://www.e-stat.go.jp/stat-search?page=1&toukei=00200522&survey=%E7%A9%BA%E3%81%8D%E5%AE%B6>

(3) 総務省：“平成25年住宅・土地統計調査”

<https://www.e-stat.go.jp/stat-search/file-download?statInfId=000028248852&fileKind=0>

(4) 総務省：“平成10年住宅・土地統計調査”

<https://www.e-stat.go.jp/stat-search/file-download?statInfId=000000050015&fileKind=0>

(5) NPO法人空き家空き地センター：“増え続ける空き家～2つの空き家問題～”

<https://www.akiya-akichi.or.jp/what/troubles/>