

# 機械学習による 要介護認定割合・平均介護費用の 社会的決定要因の分析

谷口 豊	ジブラルタ生命保険株式会社
稲垣 光隆	早稲田大学大学院会計研究科
大和田 孝文	早稲田大学大学院商学研究科

日本アクチュアリー会2019年度第1回研究集会 2019/10/26

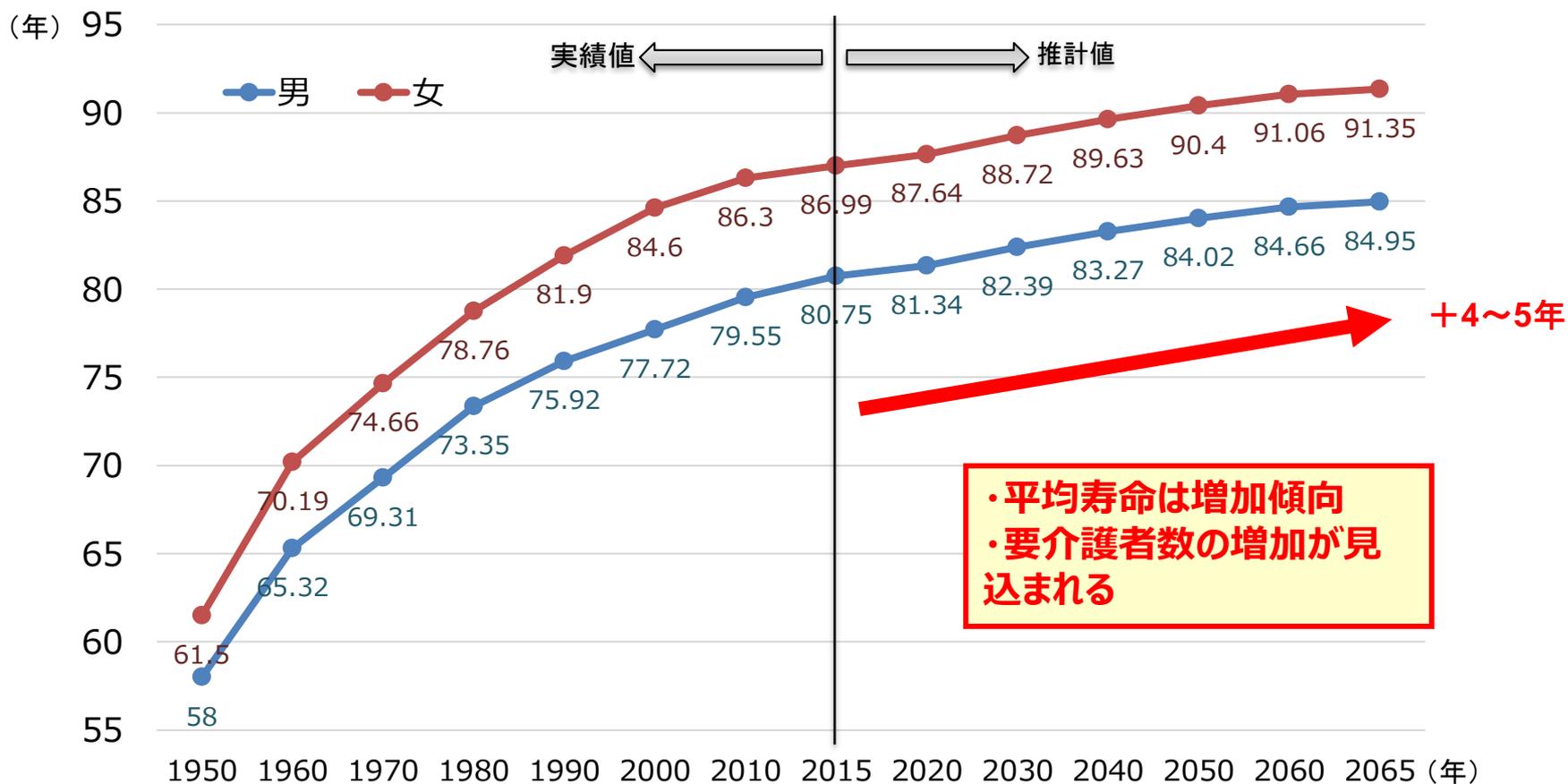
本稿の内容はすべて筆者の個人的な見解であり、筆者が所属する会社の見解とは無関係である。

- **動機・背景**
- **先行研究および本稿の新規性と貢献**
- **使用データ**
- **分析手法**
- **結果の考察**
- **まとめと課題**

## ✓ 公的介護保険は市区町村ごとの運営 運営は持続可能か？

- 市区町村ごとのソーシャルキャピタルの特性が「要介護認定割合」「平均介護費用」に与える影響を分析し、各市区町村が介護予防・介護費抑制に関し、より適切な資源配分を行うための施策を提言したい

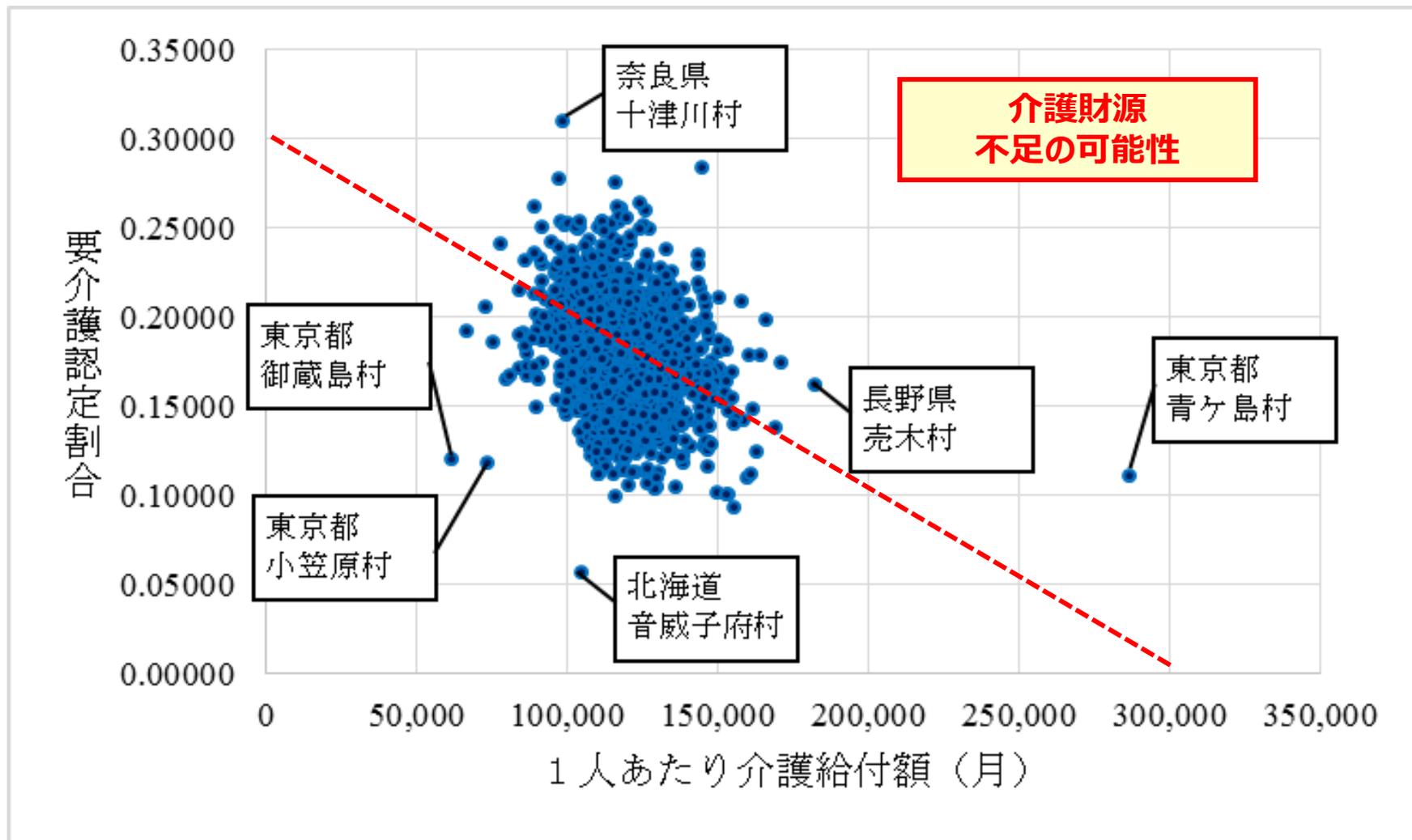
## 平均寿命の推移と将来推計



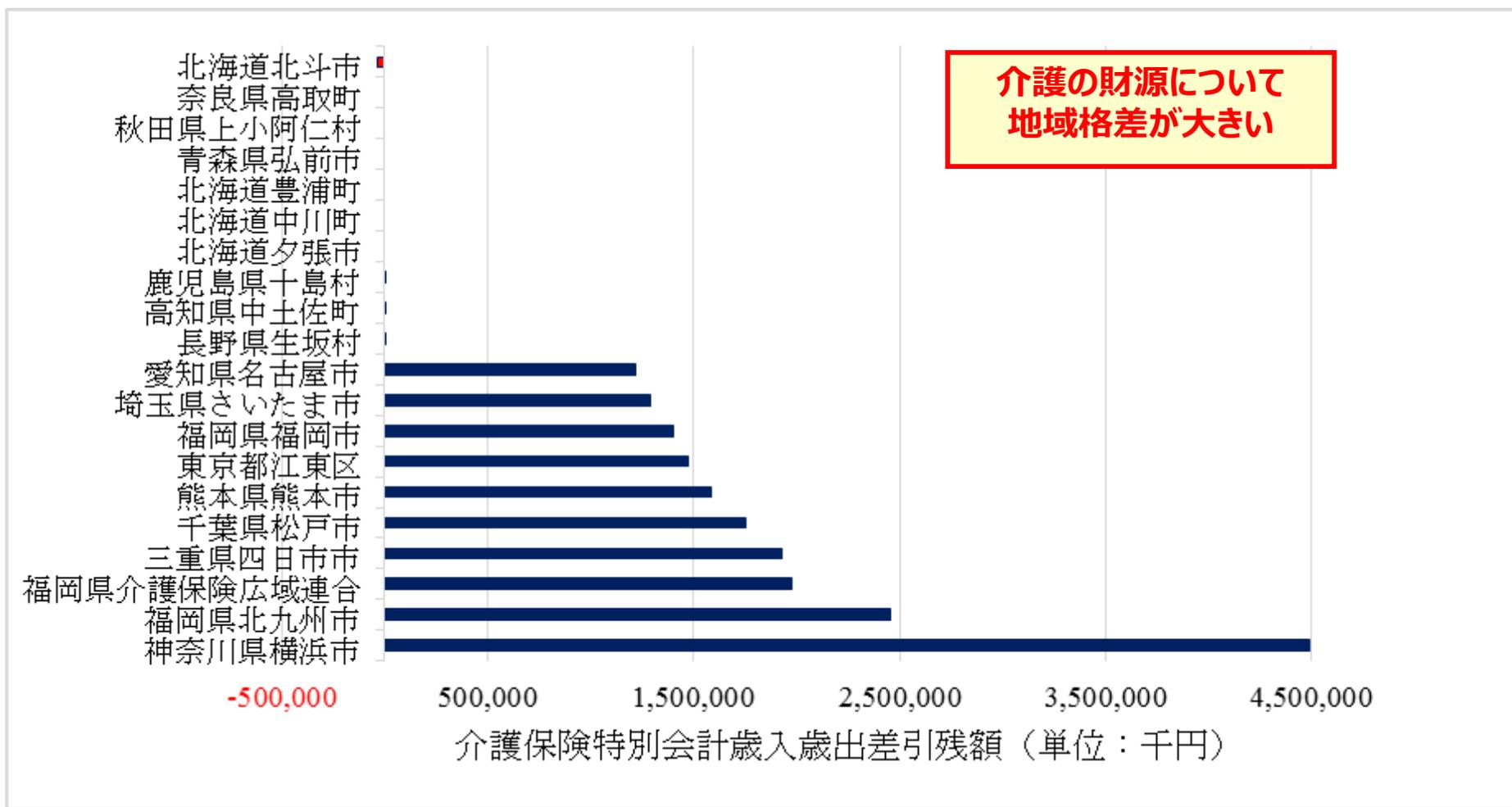
資料：1950年は厚生労働省「簡易生命表」、1960年から2015年までは厚生労働省「完全生命表」、2020年以降は国立社会保障・人口問題研究所「日本の将来推計人口（平成29年推計）」の出生中位・死亡中位仮定による推計結果をもとに筆者作成

（注）1970年以前は沖縄県を除く値である。

## 要介護認定割合と介護費用（市区町村別）



## 介護保険特別会計（市区町村別）



## 【要介護認定割合・平均介護費用の決定要因に関する研究】

- ・単相関、固定効果モデル及びOLS推定による分析を市区町村単位で実施し、予防事業への参加が、軽度要介護認定の上昇抑制効果があることを確認した  
内閣府（2018）
- ・地域在住高齢者9,702人を3年間追跡し、要介護認定のリスク要因について、服薬数が多い、咀嚼力が低い、歩行時間30分未満、友人と会う頻度月1回未満などを特定した  
平井ほか（2009）
- ・サービス種類・サービス利用量の選択に着目し、市区町村単位の回帰分析を実施  
居住サービスは、低所得者が多い地域や、高齢者夫婦世帯が多い地域で低く  
サービス利用量は、医師密度が正で有意、第一次産業が負で有意を確認した  
久保寺（2013）
- ・要介護認定割合に関し、所得段階1は正に有意であることを確認した  
安藤（2007）  
近藤他（2012）、齋藤（2018）、佐藤（2016）など

## 【機械学習を活用した研究】

- ・ランダムフォレスト、回帰木により説明変数を特定し、1人当たり医療費の都道府県のグループ分けを行った **須田（2010）**
- ・企業のデフォルト予測モデルについて、機械学習モデルとロジットモデルの精度を比較した **三浦・井實・竹川（2019）**

# 新規性

- ✓ 先行研究は、  
回帰分析等の手法であり、関連性が自明でない幅広いデータに  
関して、**機械学習を用いて分析したものでない**  
  
具体的には、  
介護供給体制を示す要因（介護老人福祉施設数など）  
市区町村の財力を示す要因（老人福祉費など）  
産業特性を示す要因（第一次産業従事者数など）  
世帯構成を示す要因（高齢単身世帯割合など）は分析対象だが  
**治療機会を示す要因**（一般診療所数など）  
**生活の豊かさを示す要因**（所得など ※特定地域については先行研究あり）  
**生活の利便性を示す要因**（公民館数・小売店数など）  
**生活衛生を示す要因**（非水洗化人口数など）は分析対象でなく、  
社会的決定要因としての説明変数が限定的である

- ✓ **特定地域について追跡調査したものであり、  
全国を調査したものではないことが多い**
- ✓ **所得に関して、低所得層の要介護リスクが高いことは分析しているが  
所得格差の要介護リスクへの影響を分析したものはない**

## 要介護認定割合

### 先行研究と統合的な結果

正の相関	低所得者数、医療・福祉産業従事者数
負の相関	農業・林業従事者数、第二次産業従事者数

### 先行研究では得られていない結果

正の相関	所得格差、一般診療所数、非水洗化人口
負の相関	肺がん検診受診者数、生活サービス・娯楽業従事者数、第二次産業、転入者数

## 平均介護費用

### 先行研究と統合的な結果

正の相関	介護老人福祉施設数
負の相関	低所得者数、高齢者夫婦世帯割合

### 先行研究では得られていない結果

正の相関	国民年金割合、第二次産業従事者数、完全失業者数、保育所数
負の相関	所得格差、一般診療所数、小学校教員数

# 使用データ

## 公的統計データの集計値（2015年・2010年の2年分）

介護保険事業状況報告（年報）	厚生労働省
介護給付費等実態調査（月報10月度）	厚生労働省
市区町村別生命表	厚生労働省
地域保健・健康増進事業報告	厚生労働省
市町村別状況	厚生労働省
国勢調査	総務省統計局
統計でみる市区町村のすがた	総務省統計局
市町村別決算状況調	総務省

### 目的変数

『介護保険事業状況報告』をもとに作成した要介護認定割合  
『介護給付費等実態調査』をもとに作成した平均介護費用  
それぞれ、65～75歳未満/75歳以上 の2区分

（注）平均介護費用は、各要介護度の平均介護費用を各要介護度数で加重平均したもの

### 説明変数

市区町村別に把握可能な統計値を幅広く候補とした  
ただし、欠損値を含むデータは、原則レコードごと削除した

→ **説明変数の候補数は124**

（注）「所得平均」は公的介護保険料の所得区分で、所得第1段階を1、所得第2段階を2、…  
とし各第一号被保険者数により加重平均したものである

（注）「所得ジニ係数」は、市区町村のうち最も人口の多い横浜市の所得基準を  
すべての市区町村にも適用し算定した

# 使用データ

項目	目的変数	説明変数	欠損値	出典	表番号
要介護（要支援）認定者数男女計（65歳以上）	○			介護保険事業状況報告 （厚生労働省）	第4-1-1表
要介護（要支援）認定者数男女計（65歳以上75歳未満）	○				第4-1-1表
要介護（要支援）認定者数男女計（75歳以上）	○				第4-1-1表
介護費用（65歳以上）	○			介護給付費等実態調査 （厚生労働省）	第6表、第7表
介護費用（65歳以上75歳未満）	○				第6表、第7表
介護費用（75歳以上）	○				第6表、第7表
所得段階(2015年：最低1と最高9、2010年：最低1と最高6)		○		介護保険事業状況報告 （厚生労働省）	第17表
所得平均		○			第17表
所得ジニ係数		○			第17表
平均余命（男性）		○		市区町村別生命表 （厚生労働省）	統計表1
平均余命（女性）		○			
外国人人口（対人口）		○		統計でみる 市区町村のすがた （総務省統計局）	
出生数（対人口）		○			
死亡数（対人口）		○			
転入者数（対人口）		○			
転出者数（対人口）		○			
昼間人口（対人口）		○			
核家族世帯数（対世帯数）		○			
単独世帯数（対世帯数）		○			
65歳以上の世帯員がいる核家族世帯数（対世帯数）		○			
高齢夫婦世帯数（対世帯数）		○			
高齢単身世帯数（対世帯数）		○			
婚姻件数（対人口）		○			
離婚件数（対人口）		○			

# 使用データ

項目	目的変数	説明変数	欠損値	出典	表番号
総面積(北方地域及び竹島を除く)(対人口)		○		統計でみる 市区町村のすがた (総務省統計局)	
可住地面積(対人口)		○			
課税対象所得(対人口)		○			
納税義務者数(所得割)(対人口)		○			
第2次産業事業所数(対人口)		○			
第3次産業事業所数(対人口)		○			
第2次産業従業者数(対人口)		○			
第3次産業従業者数(対人口)		○			
耕地面積(対人口)			23		
製造品出荷額等(対人口)			75		
製造業従業者数(対人口)		○			
商業年間商品販売額(対人口)			4		
商業事業所数(対人口)		○			
商業従業者数(対人口)		○			
財政力指数(市町村財政)			23		
実質収支比率(市町村財政)			23		
実質公債費比率(市町村財政)		○			
歳入決算総額(市町村財政)(対人口)		○			
歳出決算総額(市町村財政)(対人口)		○			
地方税(市町村財政)(対人口)		○			
幼稚園数(対人口)		○			
幼稚園在園者数(対人口)		○			
小学校数(対人口)		○			
小学校教員数(対人口)		○			
小学校児童数(対人口)		○			

(注) 欠損値の数は2015年度データのもの

# 使用データ

項目	目的変数	説明変数	欠損値	出典	表番号
中学校数（対人口）		○		統計でみる 市区町村のすがた (総務省統計局)	
中学校教員数（対人口）		○			
中学校生徒数（対人口）		○			
高等学校数（対人口）		○			
高等学校生徒数（対人口）		○			
完全失業者数（対人口）		○			
第1次産業就業者数（対人口）		○			
第2次産業就業者数（対人口）		○			
第3次産業就業者数（対人口）		○			
役員数（対人口）		○			
公民館数（対人口）		○			
図書館数（対人口）		○			
持ち家数（対人口）			527		
借家数（対人口）			527		
1住宅当たり延べ面積			527		
非水洗化人口（対人口）		○			
ごみ計画収集人口（対人口）		○			
ごみ総排出量（対人口）			23		
ごみのリサイクル率			23		
小売店数（対人口）		○	1		
飲食店数（対人口）		○	1		
大型小売店数（対人口）		○	1		
百貨店、総合スーパー数（対人口）		○	1		

(注) 欠損値の数は2015年度データのもの

# 使用データ

項目	目的変数	説明変数	欠損値	出典	表番号
一般病院数（対人口）		○		統計でみる 市区町村のすがた (総務省統計局)	
一般診療所数（対人口）		○			
歯科診療所数（対人口）		○			
医師数（対人口）		○			
歯科医師数（対人口）		○			
薬剤師数（対人口）		○			
介護老人福祉施設数（対人口）		○			
児童福祉施設等数（助産施設・児童遊園を除く）（対人口）		○			
保育所等数（対人口）		○			
保育所等在所児数（対人口）		○			
国民健康保険被保険者数（対人口）			7		
職業別就業者数_管理的職業従事者（対人口）		○		国勢調査 (総務省統計局)	
職業別就業者数_専門的・技術的職業従事者（対人口）		○			
職業別就業者数_事務従事者（対人口）		○			
職業別就業者数_販売従事者（対人口）		○			
職業別就業者数_サービス職業従事者（対人口）		○			
職業別就業者数_保安職業従事者（対人口）		○			
職業別就業者数_農林漁業従事者（対人口）		○			
職業別就業者数_生産工程従事者（対人口）		○			
職業別就業者数_輸送・機械運転従事者（対人口）		○			
職業別就業者数_建設・採掘従事者（対人口）		○			
職業別就業者数_運搬・清掃・包装等従事者（対人口）		○			
職業別就業者数_分類不能の職業（対人口）		○			

(注) 欠損値の数は2015年度データのもの

# 使用データ

項目	目的変数	説明変数	欠損値	出典	表番号
産業別就業者数_農業, 林業 (対人口)		○		国勢調査 (総務省統計局)	
産業別就業者数_漁業 (対人口)		○			
産業別就業者数_鉱業, 採石業, 砂利採取業 (対人口)		○			
産業別就業者数_建設業 (対人口)		○			
産業別就業者数_製造業 (対人口)		○			
産業別就業者数_電気・ガス・熱供給・水道業 (対人口)		○			
産業別就業者数_情報通信業 (対人口)		○			
産業別就業者数_運輸業, 郵便業 (対人口)		○			
産業別就業者数_卸売業, 小売業 (対人口)		○			
産業別就業者数_金融業, 保険業 (対人口)		○			
産業別就業者数_不動産業, 物品賃貸業 (対人口)		○			
産業別就業者数_学術研究, 専門・技術サービス業 (対人口)		○			
産業別就業者数_宿泊業, 飲食サービス業 (対人口)		○			
産業別就業者数_生活関連サービス業, 娯楽業 (対人口)		○			
産業別就業者数_教育, 学習支援業 (対人口)		○			
産業別就業者数_医療, 福祉 (対人口)		○			
産業別就業者数_複合サービス事業 (対人口)		○			
産業別就業者数_サービス業 (対人口)		○			
産業別就業者数_公務 (対人口)		○			
産業別就業者数_分類不能の産業 (対人口)		○			
健康増進に関する会議の開催回数 (対人口)		○		地域保健・健康増進事業報告 (厚生労働省)	第35表
健康増進に関する会議の参加機関・団体数 (対人口)		○		地域保健編	第35表

# 使用データ

項目	目的変数	説明変数	欠損値	出典	表番号
集団健康教育の開催回数（対人口）		○		地域保健・健康増進事業 報告（厚生労働省） 健康増進編	第04表
集団健康教育の参加延人員（対人口）		○			第04表
健康診査受診者数（65歳以上）（対65歳以上人口）		○			第0601表
健康診査受診者数 血圧（65歳以上）（対65歳以上人口）		○			第0701表
健康診査受診者数 脂質異常（65歳以上）（対65歳以上人口）		○			第0701表
健康診査受診者数 糖尿病（65歳以上）（対65歳以上人口）		○			第0701表
健康診査受診者数 貧血（65歳以上）（対65歳以上人口）		○			第0701表
健康診査受診者数 肝疾患（65歳以上）（対65歳以上人口）		○			第0701表
健康診査受診者数 腎機能障害（65歳以上）（対65歳以上人口）		○			第0701表
健康診査受診者数 たばこ（65歳以上）（対65歳以上人口）		○			第0701表
歯周疾患検診受診者数（60歳以上）（対65歳以上人口）		○			第10表
骨粗鬆症検診受診者数（65歳以上）（対65歳以上人口）		○			第11表
訪問指導従事者延人員 保健師（対人口）		○			第15表
訪問指導従事者延人員 管理栄養士及び栄養士（対人口）		○			第15表
胃がん検診受診者数 胃部エックス線検査（65歳以上） （対65歳以上人口）		○		第1601表	
肺がん検診受診者数 胸部エックス線検査（65歳以上） （対65歳以上人口）		○		第1701表	
大腸がん検診受診者数（65歳以上）（対65歳以上人口）		○		第1801表	
国民年金割合（対年金受給者数）		○		市町村別状況 （厚生労働省）	
厚生年金割合（対年金受給者数）		○			
地方債（対人口）		○		市町村別決算状況調 （総務省）	
社会福祉費（対人口）		○			
老人福祉費（対人口）		○			
保健衛生費（対人口）		○			

## ステップ1： 機械学習による説明変数の一次選択

ランダムフォレスト（%IncMSE、IncNodePurity）、XGBoost（xgbTree）を実行し、指標重要度の高い順から30位までを1カウント（Rfは各0.5カウント）合計カウント数で上位20%までを一次選択する



## ステップ2： 重回帰分析による説明変数の特定

一次選択された説明変数について、さらに重回帰分析により選択を行う  
重回帰分析はステップワイズの手法を用いる



## ステップ3： Elastic Netによる妥当性検証

ステップワイズで特定した説明変数について、Elastic Netで得られる結果と照合を行い、その妥当性を確認する

## ステップ1： 機械学習による説明変数の一次選択

モデル		特徴
ランダムフォレスト	%IncMSE	ブートストラップ法でランダムに抽出されたデータにより作成した回帰木に対し、抽出されなかったデータを当てはめた結果より算出される特徴量である。試行処理のため計算時間を要する。
	IncNodePurity	枝を分岐させる際の平均二乗誤差により算出された特徴量である。計算時間は短いが精度は劣る。
XGBoost	xgbTree	決定木を構築する際、それまでに作られたすべての決定木の予測結果を利用し予測できなかった誤差を新たに目的変数とし、逐次的に決定木を構築していくモデルである。逐次的な構築のアルゴリズムに非線形モデルを用いる。

ランダムフォレストは過学習しにくいアルゴリズムであるが、XGBoostは過学習しやすい  
→ クロスバリデーション等を用いてパラメータチューニングする必要がある

XGBoostのパラメータチューニングは、パッケージcaretを用いた。  
データを10分割してクロスバリデーションを行い、  
各パラメータ4通りの組み合わせについて二乗平均平方根誤差（RMSE）を求め  
RMSEが最小値となるパラメータを採用した。

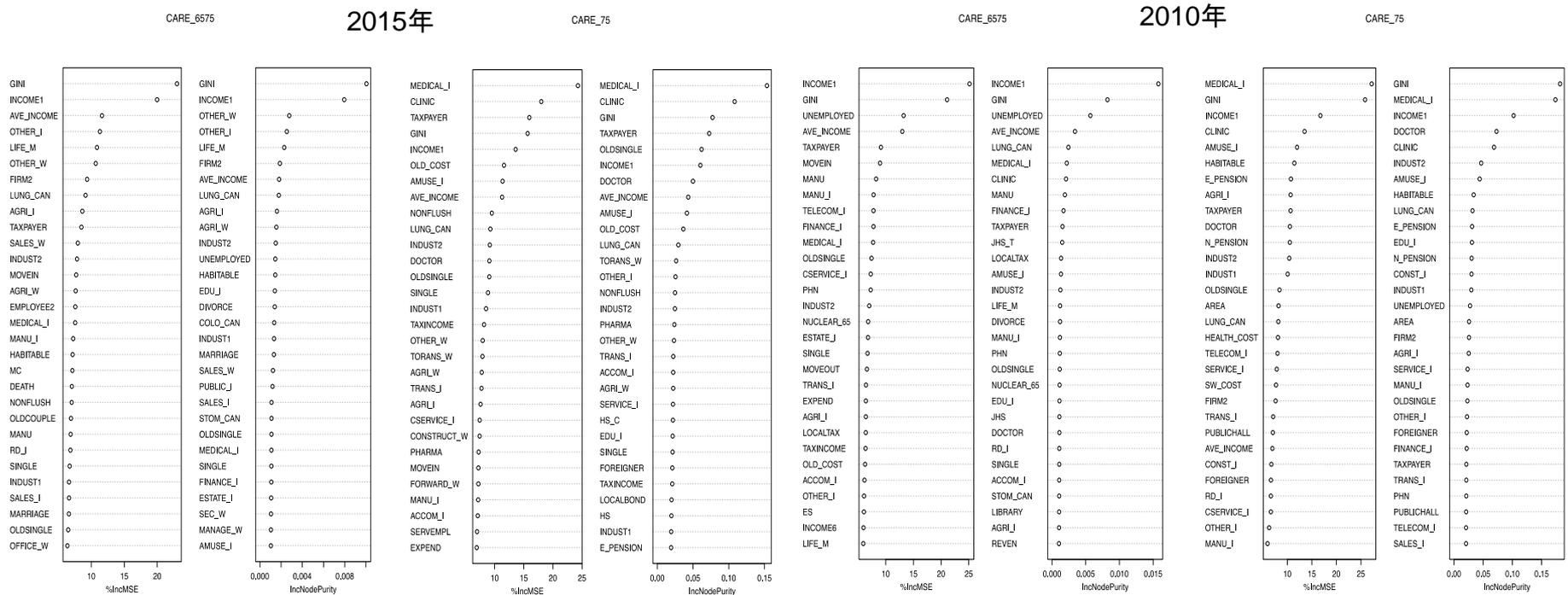
## ステップ<sup>o</sup>2 : 重回帰分析による説明変数の特定

## ステップ<sup>o</sup>3 : Elastic Netによる妥当性検証

モデル	特徴
<b>ステップワイズ法</b>	<p>統計学モデルにおける説明変数の選択手法の一つであり、変数を1ずつ追加・削除することにより評価関数（本稿ではAIC）を逐次計算し、最適な説明変数の組み合わせを探索する手法である。</p> <p>(注) 多重共線性の有無についてはVIFを用いて評価し、相関が認められる説明変数については単回帰分析を行い決定係数の高い方を残した。</p>
<b>ElasticNet</b>	<p>Ridge回帰、Lasso回帰の双方の問題を解消したアルゴリズムである。Ridge回帰は、多重共線性の問題を解消可能であるが、特定の説明変数のみを抽出できず回帰係数全体が縮小推定される問題がある。Lasso回帰は、この問題を解消したモデルであるもののGrouping効果を考慮できない等の問題がある。ElasticNetはこれらの問題を解消したアルゴリズムである。</p>

## ステップ1：機械学習による説明変数の一次選択の結果

### ランダムフォレスト【要介護認定割合】



- ・重要度の高い説明変数は、所得ジニ係数、所得段階1、第二次産業従事者数、平均所得、医療・福祉系産業従事者数となった。
- ・65歳以上75歳未満は、総じて所得に関連した説明変数の重要度が高い。
- ・75歳以上は、医療サービスに関連のある説明変数の重要度が高い。

## ランダムフォレスト【平均介護費用】

CARECOST\_6575

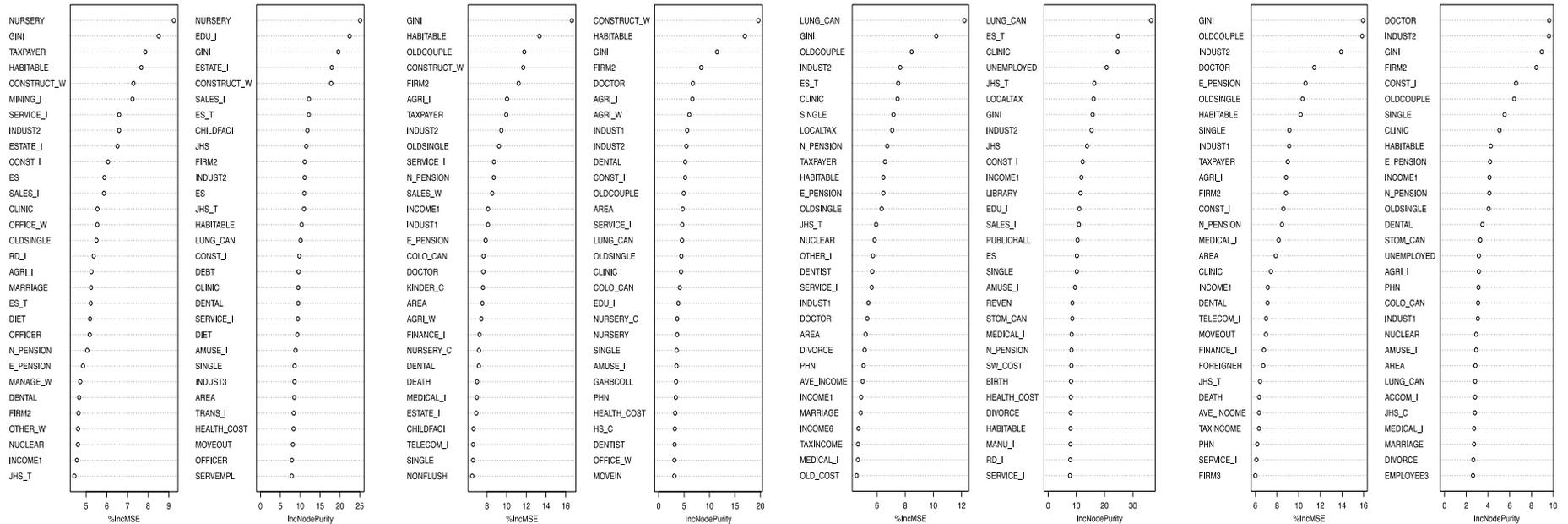
2015年

CARECOST\_75

CARECOST\_6575

2010年

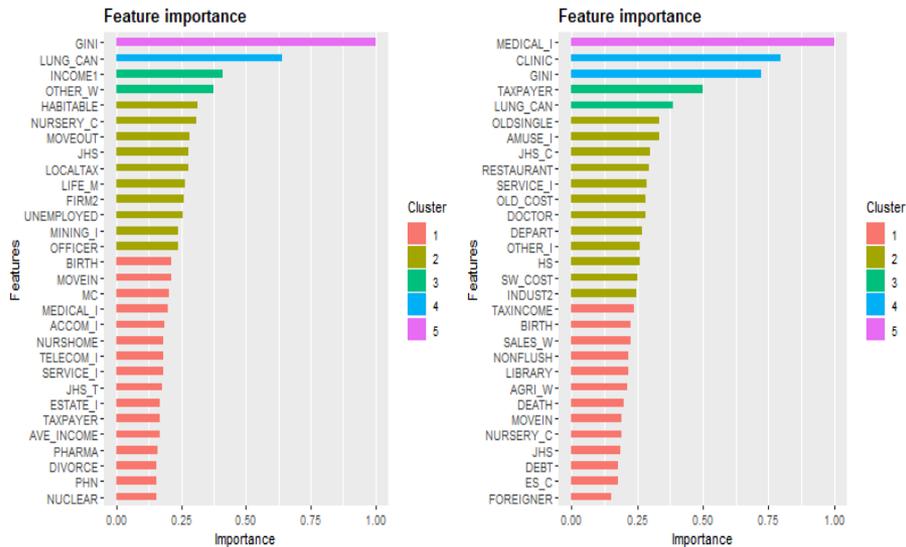
CARECOST\_75



- 重要度の高い説明変数は、所得ジニ係数、第二次産業従事者数、肺がん検診受診者数、可住地面積、一般診療所数となった。
- 65歳以上75歳未満は、2015年度では、保育所数が最も重要度の高い説明変数であるが、2010年度の分析結果では肺がん検診受診者数が最も重要度の高い説明変数である。

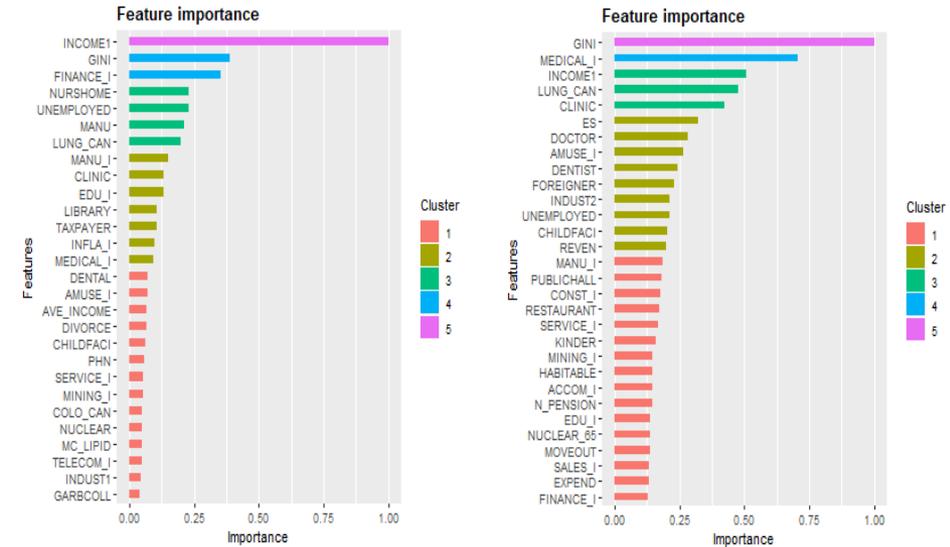
## XGBoost 【要介護認定割合】

2015年



(注) 左：65歳以上75歳未満 右：75歳以上

2010年



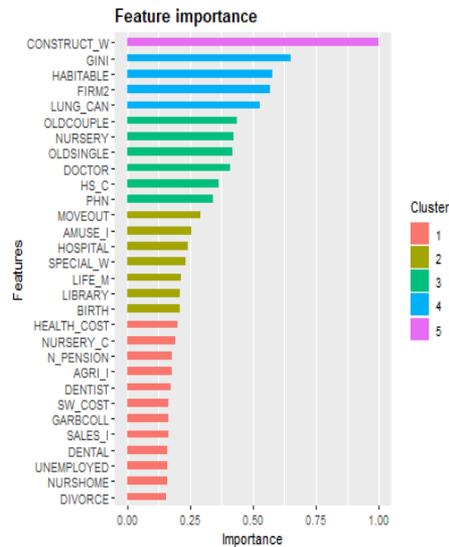
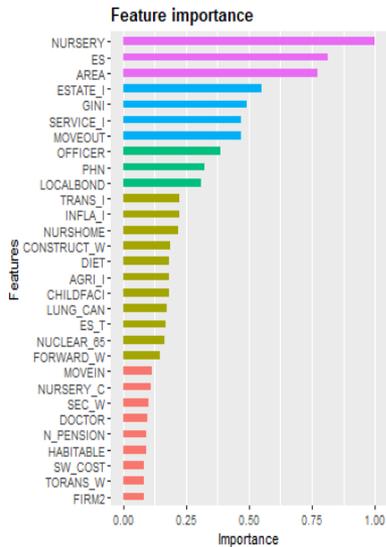
(注) 左：65歳以上75歳未満 右：75歳以上

- 重要度の高い説明変数は、所得ジニ係数、肺がん検診受診者数、医療・福祉系産業従事者数、生活関連サービス業・娯楽産業従事者数、一般診療所数となった。
- 所得に関する説明変数の重要度の傾向は、ランダムフォレストの結果と同一。
- 75歳以上で、医療サービスに関連の重要度が高いという結果も、ランダムフォレストと共通。

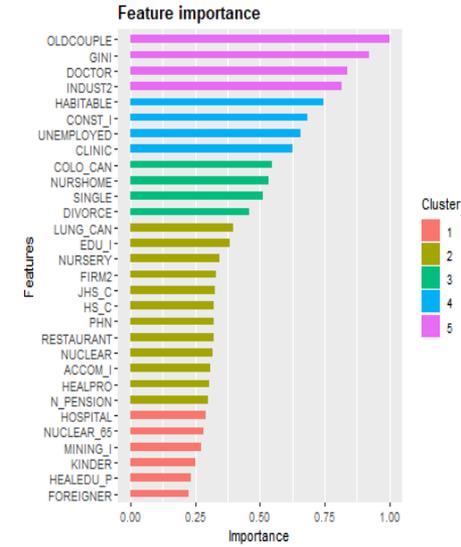
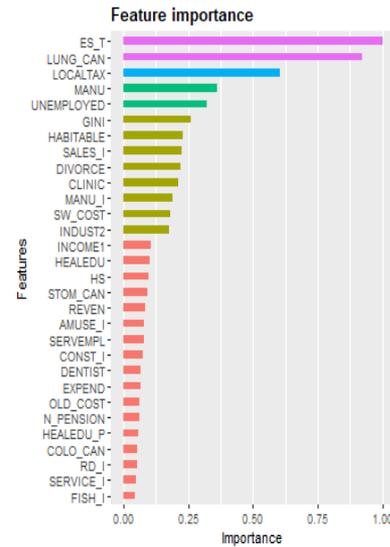
## XGBoost 【平均介護費用】

2015年

2010年



(注) 左：65歳以上75歳未満 右：75歳以上



(注) 左：65歳以上75歳未満 右：75歳以上

- 重要度の高い説明変数は、所得ジニ係数、可住地面積、国民年金割合、保育所数、介護老人福祉施設数、肺がん検診受診者数となった。
- 肺がん検診受診者数の重要度が高い傾向は、ランダムフォレストと同様の結果となった。
- 保育所数や、介護老人福祉施設数などの重要度が高い傾向は、ランダムフォレストより顕著である。

# 結果の考察

## 機械学習の結果

## → 重回帰分析 (step wise)

No	要介護認定割合		介護平均費用	
	一次選択変数	件数	一次選択変数	件数
1	所得ジニ係数	8	所得ジニ係数	8
2	産業別就業者数 医療・福祉	7.5	可住地面積	7.5
3	所得段階1	7	肺がん検診受診者数 胸部エックス線検査 (65歳以上)	7.5
4	肺がん検診受診者数 胸部エックス線検査 (65歳以上)	7	国民年金割合	6.5
5	産業別就業者数 生活関連サービス業・娯楽業	6.5	第二次産業就業者数	6
6	産業別就業者数 サービス業	6.5	医師数	5.5
7	納税義務者数 (所得割)	6.5	第二次産業事業所数	5.5
8	第二次産業就業者数	6	産業別就業者数 農業・林業	5
9	所得平均	5.5	一般診療所数	5
10	一般診療所数	5.5	保育所等数	5
11	完全失業者数	5	訪問指導従事者延人員 保健師	5
12	高齢単身世帯数	4.5	産業別就業者数 サービス業	5
13	産業別就業者数 宿泊業・飲食サービス業	4	総面積	4.5
14	医師数	4	産業別就業者数 建設業	4.5
15	産業別就業者数 教育・学習支援業	4	離婚件数	4.5
16	外国人人口	4	高齢夫婦世帯数	4.5
17	可住地面積	4	職業別就業者数 建設・採掘従事者	4
18	(除外)第1次産業就業者数	4	小学校教員数	4
19	産業別就業者数 分類不能の産業	4	介護老人福祉施設数	4
20	産業別就業者数 金融業・保険業	3.5	高齢単身世帯数	4
21	第二次産業事業所数	3.5	産業別就業者数 卸売業, 小売業	4
22	(除外)製造業	3.5	単独世帯数	4
23	転入者数	3.5	社会福祉費	4
24	非水洗化人口	3.5	完全失業者数	4
25	産業別就業者数 農業・林業	3	産業別就業者数 生活関連サービス業, 娯楽	3.5
26	職業別就業者数 農林漁業従事者	3	歯科診療所数	3.5
27	離婚件数	3	所得段階1	3.5
28	図書館数	3		
29	平均余命 (男性)	3		
30	産業別就業者数 鉱業・採石業・砂利採取業	3		
31	職業別就業者数 分類不能の職業	3		
32	産業別就業者数 情報通信業	3		

# 結果の考察

## ステップ2：重回帰分析（step wise）による説明変数の特定

説明変数	要介護認定割合	
	65-75歳未満	75歳以上
切片	-0.0017	-0.0024
所得段階1	0.1905 ***	0.1305 ***
肺がん検診受診者数 胸部エックス線検査（65歳以上）	-0.0485 **	-0.0445 **
産業別就業者数 分類不能の産業	0.1698 ***	0.1158 ***
産業別就業者数 サービス業	0.0830 ***	0.0399 *
産業別就業者数 医療・福祉	0.1149 ***	0.2643 ***
産業別就業者数 教育・学習支援業	-0.0765 ***	-
産業別就業者数 生活関連サービス業・娯楽業	-0.0702 ***	-0.1051 ***
産業別就業者数 宿泊業・飲食サービス業	-0.0536 **	-0.0427 *
産業別就業者数 金融業・保険業	0.1120 ***	0.1439 ***
産業別就業者数 情報通信業	-0.0878 **	-0.1664 ***
産業別就業者数 鉱業・採石業・砂利採取業	-0.0229	-0.0726 ***
産業別就業者数 農業・林業	-0.1154 ***	-0.1046 ***
一般診療所数	-	0.0968 ***
非水洗化人口	0.0588 *	0.0921 ***
図書館数	-0.0748 ***	0.0541 ***
第二次産業就業者数	-0.0691 **	-0.1175 ***
完全失業者数	0.0528 **	-
納税義務者数（所得割）	-0.0998 **	0.0980 **
可住地面積	0.0530 *	-0.0458 .
高齢単身世帯数	-0.0611 *	-
転入者数	-0.1376 ***	-0.1236 ***
外国人人口	0.0345 *	-
平均余命（男性）	-0.1046 ***	-
所得ジニ係数	0.1926 ***	0.1338 ***
Residual standard error:	0.8158	0.8477
Multiple R-squared:	0.3381	0.2838
Adjusted R-squared:	0.3332	0.2795
F-statistic:	69.4565	65.3230
p-value	2.20E-16	2.20E-16

正の相関・・・

所得ジニ係数、  
所得段階1、  
医療・福祉産業、  
分類不能産業、  
金融業・保険業、  
非水洗化人口

負の相関・・・

生活関連サービス業・娯楽産業、  
農業・林業、  
第二次産業、  
情報通信業、  
転入者数

# 結果の考察

説明変数	平均介護費用	
	65-75歳未満	75歳以上
切片	-0.0013	0.0015
所得段階1	-0.0675 **	-
社会福祉費	0.0767 **	0.0530 .
国民年金割合	0.1847 ***	0.1649 ***
肺がん検診受診者数 胸部エックス線検査 (65歳以上)	-	0.0507 **
訪問指導従事者延人員 保健師	-	0.0758 **
産業別就業者数 サービス業	0.0511 **	0.0554 **
産業別就業者数 生活関連サービス業・娯楽業	-	0.0798 ***
産業別就業者数 卸売業・小売業	-	-0.0700 **
産業別就業者数 建設業	-	0.0960 ***
産業別就業者数 農業・林業	-	0.0584 *
保育所等数	0.1798 ***	0.1438 ***
介護老人福祉施設数	0.0886 ***	0.1007 ***
医師数	0.0434 *	-
歯科診療所数	-	0.0429 *
一般診療所数	-0.1724 ***	-0.0770 ***
第二次産業就業者数	0.0734 **	0.0933 ***
完全失業者数	0.0791 ***	0.0694 ***
小学校教員数	-0.2457 ***	-0.0750 *
第二次産業事業所数	0.0339	0.0765 ***
可住地面積	-	-0.0519 .
離婚件数	0.0414	0.0363
高齢単身世帯数	0.0657 .	0.0826 *
高齢夫婦世帯数	-0.0813 **	-0.2072 ***
単身世帯数	-	-0.0549 .
所得ジニ係数	-0.2003 ***	-0.2399 ***
Residual standard error:	0.8983	0.8590
Multiple R-squared:	0.1969	0.2677
Adjusted R-squared:	0.1928	0.2623
F-statistic:	48.0365	49.7082
p-value	2.20E-16	2.20E-16

## 正の相関・・・

国民年金割合、  
保育所数、  
介護老人福祉施設数、  
第二次産業、  
完全失業者数

## 負の相関・・・

一般診療所数、  
高齢夫婦世帯数、  
所得ジニ係数

## ステップ3： Elastic Netによる妥当性検証

### 一次選択した説明変数を用いてElastic Netを実施

ステップワイズで最終的に選ばれた説明変数と比較

#### 要介護認定割合

65歳以上75歳未満の適合率は 85.9%

75歳以上の適合率は 79.7%

#### 平均介護費用

65歳以上75歳未満の適合率は 56.3%

75歳以上の適合率は 88.9%

→ 平均介護費用の65歳以上75歳未満での適合率は高いとは言えないが、その他の適合率は概ね高く、ステップワイズで選ばれた説明変数はElastic Netの結果とほぼ等しいと言える。

※適合率は

$$\frac{(\text{ステップワイズで選択された説明変数がElastic Netでも選択される数})}{(\text{ステップワイズで選択された説明変数の総数})} \text{ と } \frac{(\text{Elastic Netで選択された説明変数がステップワイズでも選択される数})}{(\text{Elastic Netで選択された説明変数の総数})}$$
の平均

# まとめと課題

## 結果

ソーシャルキャピタル	説明変数	先行研究の有無および内容		本稿の結果
治療機会	一般診療所	×		平均介護費用に関し負の相関
介護予防	肺がん検診	×		要介護認定割合に関し負の相関
市区町村の財力		○	要介護認定割合に関し自治体の財政力指数は負の相関	
介護供給体制	介護老人福祉施設	○	1人当たりの介護サービス費用に関し介護供給体制は正の相関	平均介護費用に関し正の相関
産業特性	農業・林業	○	1人当たりサービス利用件数について第一次産業比率は負の相関 要介護認定割合に関し農林業従事者数は負の相関	要介護認定割合に関し負の相関
	第二次産業	○/×	要介護認定割合に関し第二次産業従事者数は負の相関	要介護認定割合に関し負の相関 平均介護費用に関し正の相関
	医療・福祉産業	○	1人当たりサービス利用件数について医師密度は正の相関	要介護認定割合に関し正の相関
	金融業・保険業	×		要介護認定割合に関し正の相関
	情報通信業	×		要介護認定割合に関し負の相関
	生活関連サービス・娯楽業	×		要介護認定割合に関し負の相関
	宿泊業・飲食サービス業	×		要介護認定割合に関し負の相関
世帯構成	高齢者単身世帯	○	高齢者単身世帯割合は居住サービスへのニーズが高い	
	高齢者夫婦世帯	○	高齢者夫婦世帯割合は居住サービスへのニーズが低い	平均介護費用に関し負の相関
人口動態	転入者数	×		要介護認定割合に関し負の相関
生活の豊かさ	所得段階1	○	要介護認定割合に関し所得段階1は正に有意 低所得者が多い地域では居住サービスへのニーズが低い	要介護認定割合に関し正の相関
	所得ジニ係数	×		要介護認定割合に関し正の相関 平均介護費用に関し負の相関
	完全失業者	×		平均介護費用に関し正の相関
生活の利便性	保育所	×		平均介護費用に関し正の相関
	小学校教員	×		平均介護費用に関し負の相関
生活衛生	非水洗化人口	×		要介護認定割合に関し正の相関
年金種類	国民年金割合	×		平均介護費用に関し正の相関

## 先行研究と統合的な結果、および 先行研究では得られていない結果を確認した

### 先行研究と統合的な結果

- ✓ 要介護認定割合について、  
低所得者数は正の相関、農業・林業は負の相関
- ✓ 平均介護費用について、  
医療・福祉産業は正の相関、高齢夫婦世帯数は負の相関

### 先行研究では得られていない結果

- ✓ 要介護認定割合について、  
**所得格差、非水洗化人口**は正の相関  
**生活サービス・娯楽産業、肺がん検診受診者数**は負の相関
- ✓ 平均介護費用について、  
**国民年金割合、第二次産業、完全失業者数**は正の相関  
**所得格差、一般診療所数、小学校教員数**は負の相関

## 課題

- ✓ 推測について確証を得るための追加分析
- ✓ XGBoostのパラメータチューニングの充実
- ✓ 変数選択を行った上での重回帰分析のため、選択バイアスが生じている可能性を否定できない
- ✓ 要介護のソーシャルキャピタルの要因は複合的であり、複合的要素を有機的に分析するためには集計表ではなく個票データを用いる必要がある

### 今後の課題

**「所得格差」「国民年金割合」「一般診療所数」「非水洗化人口」「肺がん検診受診者数」については、今後の介護予防・介護費用抑制の施策上重要であり、データの蓄積と分析が必要**

- [1] 安藤道人 (2007) 「介護給付水準と介護保険料の地域差の実証分析－保険者データを用いた分析－」『季刊社会保障研究』Vol.44 No.1 pp.94-109
- [2] 医療科学研究所 (2014) 「健康の社会的決定要因に関する国内外の調査研究動向」2014年3月
- [3] 久保寺重行 (2013) 「介護サービス需要行動に関する実証分析－今後の介護保険制度改革に向けて－」『社会福祉学』第54巻第2号 2013年6月 pp.70-82
- [4] 近藤克則・芦田登代・平井寛・三澤仁平・鈴木佳代 (2012) 「高齢者における所得・教育年数別の死亡・要介護認定率とその性差－AGESプロジェクト縦断研究－」『医療と社会』Vol.22 No.1 2012 pp.19-30
- [5] 齋藤香里 駒村康平編著 (2018) 「介護と貧困」『貧困』ミネルヴァ書房 pp.130-142
- [6] 佐藤哲彰 (2016) 「要介護状態の発生率は、所得水準によってどう異なるか－ロジスティック回帰による分析－」『千葉商大紀要』第53巻第2号 2016年3月 pp.93-106
- [7] 須田茂夫 (2018) 「機械学習による都道府県別医療費の分析」『社会保障研究』Vol.3 No.3, 2018 (pp.403-415)
- [8] 飛田英子 (2015) 「介護費用の適正化に向けた課題－保険者データの分析を踏まえて－」『JRIレビュー』2015 Vol.11, No.30 2015年9月 pp.21-40
- [9] 内閣府政策統括官 (2018) 「要介護（要支援）認定率の地域差要因に関する分析」内閣府政策統括官（経済財政分析担当）
- [10] 西野辰哉、大森数馬 (2014) 「一中学校区を基本とする日常生活圏域設定の妥当性検討－地方中核都市における高齢者福祉行政単位と高齢者の行動実態との比較考察－」『日本建築学会計画系論文集』第79巻第699号 2014年5月 pp.1109-1118
- [11] 日本学術会議 (2011) 「わが国の健康の社会格差の現状理解とその改善に向けて」基礎医学委員会・健康・生活科学委員会合同 パブリックヘルス科学分科会 2011年9月
- [12] 平井寛、近藤克則、尾島俊之、村田千代栄 (2009) 「地域在住高齢者の要介護認定のリスク要因の検討」『日本公衆誌』第56巻第6号 2009年8月 pp.501-512
- [13] 堀口健治、弦間正彦 (2017) 「自営農業者の長寿傾向と 後期高齢者医療費への反映－埼玉県本庄市における調査を踏まえて－」『農林金融』2017年9月 pp.56-61
- [14] 三浦翔、井實康幸、竹川正浩 (2019) 「入出金情報を用いた信用リスク評価－機械学習による実証分析－」『日本銀行ワーキングペーパーシリーズ』2019年6月
- [15] Zou H, Hastie T (2005) "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society B*, 67(2) pp.301-320