

2020年からの新型コロナウイルス感染者数に関する統計解析

2018SS024 伊藤佑里恵

指導教員:白石高章

1 はじめに

世界で初めて新型コロナウイルスの症例が確認された2019年12月、日本で初めて確認された2020年1月から2年が経過している。そんな中、2021年に入ってからワクチン接種が開始されている。2021年の終わり頃には感染者数が減少傾向にあったが、新しい変異株も確認され、油断できない状況となっている。

このように2020年とは状況が大きく異なるため、2020年からの新型コロナウイルス感染者数にはどのような変化があるのかに興味を持った。そこで日本全体の新型コロナウイルス感染者数と、ワクチン接種率などの関係を分析することで、2020年とは新型コロナウイルス感染者数にどのような違いがあるのかを考察する。

2 データと分析方法

本研究では、2020年1月から2021年11月における都道府県ごとの新型コロナウイルス感染者数 ([1])、厚生労働省 ([2]) と都道府県ランキング ([3]) などのデータを調査し、本文で活用した。データは以下の通りである。 x_1 : 新型コロナウイルス感染者数, x_2 : ワクチン接種率 (1回目), x_3 : ワクチン接種率 (2回目), x_4 : 人口密度, x_5 : 平均気温, x_6 : 平均湿度, x_7 : 糖尿病による死亡者数, x_8 : 喫煙率, x_9 : 飲食店数, x_{10} : カラオケ店舗数, x_{11} : 自家用車保有台数, x_{12} : 病院数, x_{13} : GDP, x_{14} : 老年 (65歳以上) 人口率, x_{15} : 累計死亡者数, x_{16} : 公共交通機関での通勤通学率, x_{17} : 自家用車での通勤通学率, x_{18} : 重症者数。また、観測値は人口10万人あたりのものを使用する。

分析方法として、相関係数、重回帰分析、クラスター分析を行った。分析は青木 [4]、金 [5] を参考にして行った。

3 相関係数

新型コロナウイルス感染者数 (x_1) と様々なデータとの相関係数を求めることで、関連性を調べる。ここでは、各データに対して無相関検定を行い p 値が棄却されたデータに対して考察を行う。

3.1 分析結果

表1 相関係数

変数	相関係数	変数	相関係数
x_2	-0.737	x_{11}	-0.682
x_3	-0.729	x_{12}	-0.346
x_4	0.698	x_{13}	0.644
x_5	0.463	x_{14}	-0.837
x_6	-0.235	x_{15}	0.837
x_7	-0.548	x_{16}	0.690
x_8	-0.040	x_{17}	-0.714
x_9	0.318	x_{18}	0.432
x_{10}	0.231		

p 値が有意水準 5% で棄却された変数は、ワクチン接種率 (x_2, x_3), 人口密度 (x_4), 平均気温 (x_5), 糖尿病による死亡者数 (x_7), 飲食店数 (x_9), 自家用車保有台数 (x_{11}), 病院数 (x_{12}), GDP (x_{13}), 老年 (65歳以上) 人口率 (x_{14}), 累計死亡者数 (x_{15}), 公共交通機関での通勤通学率 (x_{16}), 自家用車での通勤通学率 (x_{17}), 重症者数 (x_{18}) であった。

3.2 相関係数における考察

新型コロナウイルス感染者数 (x_1) と強い正の相関を示したものに人口密度 (x_4) と GDP (x_{13}) と累計死亡者数 (x_{15}) と公共交通機関での通勤通学率 (x_{16}) があった。このことから人口が多く発展しており、主な交通手段が公共交通機関となる都心部でコロナ感染のリスクが高いのではないかと考えられる。

一方で、新型コロナウイルス感染者数 (x_1) と強い負の相関を示したものにワクチン接種率 (1回目) (x_2) とワクチン接種率 (2回目) (x_3) と自家用車保有台数 (x_{11}) と老年 (65歳以上) 人口率 (x_{14}) と自家用車での通勤通学率 (x_{17}) があった。このことから、ワクチン接種の効果はおおいにありと考える。そして、自家用車を使って移動した方が感染リスクが低いと考えられる。

4 重回帰分析

新型コロナウイルス感染者数 (x_1) を目的変数とし、ワクチン接種率 (1回目) (x_2) から重症者数 (x_{18}) を説明変数として、重回帰分析を行った。

4.1 分析結果

ステップワイズ法により、ワクチン接種率 (1回目) (x_2), 人口密度 (x_4), 糖尿病による死亡者数 (x_7), 自家用車保有台数 (x_{11}), GDP (x_{13}), 自家用車での通勤通学率 (x_{17}) が減少した。よって、表2の変数を説明変数に置いたモデルが最適となった。

表2 重回帰分析

説明変数	回帰係数	標準誤差	p 値
x_3	-2535	1543	0.110
x_5	54.31	19.84	0.001
x_6	17.85	6.487	0.001
x_8	45.60	17.73	0.015
x_9	1.169	0.438	0.012
x_{10}	23.73	17.37	0.181
x_{12}	-30.52	12.64	0.021
x_{14}	-55.55	18.04	0.004
x_{15}	35.80	6.677	0.000
x_{16}	1209	319.2	0.001
x_{18}	-1783	1144	0.128

4.2 重回帰分析における考察

決定係数は 0.9491, 自由度修正済み決定係数は 0.933 となった。

老年 (65 歳以上) 人口率 (x_{14}) と累計死亡者数 (x_{15}) が新型コロナウイルス感染者数 (x_1) に大きく影響しているのは予想できていた。公共交通機関での通勤通学率 (x_{16}) が新型コロナウイルス感染者数 (x_1) に大きく影響しており、これはテレワークや時差出勤を行う意味が大いにあると考える。また、病院数 (x_{12}) も影響を与えており、ワクチン接種を進めるだけではなく、医療施設の充実も目指す必要があると考える。

5 クラスタ分析

クラスタ分析をウォード法を用いて行った。また、データは重回帰分析におけるステップワイズ法で最適となったものを使用した。

5.1 分析結果

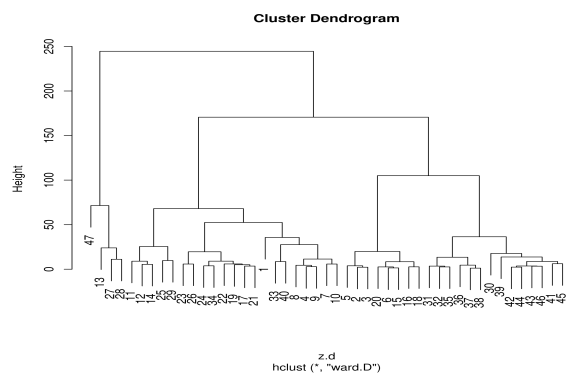


図 1 クラスタ分析

第 1 群-a:自家用車での通勤通学率が高い都道府県,
-b:GDP が高い都道府県

第 2 群:公共交通機関での通勤通学率が高い都道府県

第 3 群:自家用車数保有台数が多い都道府県

第 4 群:ワクチン接種率が高い都道府県

第 5 群:病院数が多い都道府県

5.2 クラスタ分析における考察

第 1 群:(a:沖縄 (47), b:東京 (13), 大阪 (27), 兵庫 (28))

第 1 群の a と b は共に新型コロナウイルス感染者数 (x_1) が最も多い群である。a:沖縄 (47) は自家用車保有台数 (x_{11}) が多い。観光地であり、様々な地域から観光客が訪れるため、感染者数が多いと考えられる。自家用車を積極的に利用し、この県を訪れるのは控えるべきである。

一方で、b:東京 (13), 大阪 (27), 兵庫 (28) は公共交通機関での通勤通学率 (x_{16}) が高くなっている。大都市であるため、公共交通機関を利用する人が多いと考えられる。自家用車保有台数 (x_{11}) が少なく、通勤通学以外でも公共交通機関を頻繁に利用することが多いため、感染症対策を徹底して日常生活を送る必要があると考える。

第 2 群:(埼玉 (11), 千葉 (12), 神奈川 (14), 滋賀 (25), 奈良 (29))

公共交通機関での通勤通学率 (x_{16}) が高い群である。自家用車保有台数 (x_{11}) が少ないため、公共交通機関での通勤通学率 (x_{16}) が高くなっている。この群に含まれる

都道府県は大都市に隣接しており、公共交通機関を利用して大都市へ通勤や通学をする人が多いと考えられるため、感染症対策を強化して生活する必要がある。

第 3 群:(愛知 (23), 京都 (26), 三重 (24), 広島 (34), 静岡 (22), 山梨 (19), 石川 (17), 岐阜 (21), 北海道 (1), 岡山 (33), 福岡 (40), 茨城 (8), 宮城 (4), 栃木 (9), 福島 (7), 群馬 (10))

自家用車数保有台数 (x_{11}) が最も多い群である。この群は都道府県によって飲食店数 (x_9) が多かったり、公共交通機関での通勤通学率 (x_{16}) が高かったりと様々な特徴がみられる。そのため、各都道府県で個別の感染症対策をとる必要があると考える。

第 4 群:(秋田 (5), 青森 (2), 岩手 (3), 長野 (20), 山形 (6), 新潟 (15), 富山 (16), 福井 (18))

ワクチン接種率 (x_2, x_3) が最も高い群である。そのため、新型コロナウイルス感染者数 (x_1) は最も少ない。また、自家用車での通勤通学率 (x_{17}) が高い。今後もワクチン接種をすすめる、自家用車を積極的に利用していくとよいと考える。

第 5 群:(鳥根 (31), 鳥取 (32), 山口 (35), 徳島 (36), 香川 (37), 愛媛 (38), 和歌山 (30), 高知 (39), 長崎 (42), 大分 (44), 熊本 (43), 鹿児島 (46), 佐賀 (41), 宮崎 (45))

病院数 (x_{12}) が最も多い群である。そのため、ワクチン接種率 (x_2, x_3) も比較的高い都道府県が多いが、まだ低い都道府県もあるため、今後もワクチン接種を進める必要がある。

6 おわりに

本研究を行った結果、新型コロナウイルスに対してワクチン接種の効果がおおいにあると考えることができた。しかし、病院数が多ければ必ずワクチン接種率が高いとはいえなかった。そのため、医療施設の充実だけではなく、すべての人々がいち早くワクチンを接種できる環境や制度を整える必要があると考える。また、新型コロナウイルス感染者数と通勤や通学の方法が大きく関係していることから、今後もテレワークや時差出勤、オンライン授業などを推奨していくべきであると考えられる。その他にも飲食店の数などの影響がみられたため、これ以上に新型コロナウイルス感染者数を増やさないために、都道府県ごとに必要な感染症対策を徹底していく必要があると考える。

参考文献

- [1] 総務省統計局 (<https://www.stat.go.jp/index.html>) (参照日 2022 年 1 月 17 日)
- [2] 厚生労働省 (<https://www.mhlw.go.jp/stf/covid-19/kokunainohasseijoukyou.html>) (参照日 2022 年 1 月 17 日)
- [3] 都道府県ランキング (<https://uub.jp/>) (参照日 2022 年 1 月 17 日)
- [4] 青木 繁伸:『R による統計解析』オーム社, 東京, 2009
- [5] 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』森北出版, 東京, 2007