

2020年から現在にかけての 新型コロナウイルス感染者数における統計解析

2019SS039 小林 亮生

指導教員：白石 高章

1 はじめに

一昨年、昨年から世界中で感染拡大を続ける新型コロナウイルスは現在、収束の一途を辿りつつも、未だに人類の脅威となっている。対策として2021年度には第一回、第二回の新型コロナウイルスに対するワクチン接種が実施されており、現在では三回目の摂取を勧める方針が取られている。私自身の周りにも感染者やワクチン接種済みの人間が多く、そこからこのような世界を脅かすウイルスがどのような変化を歩んできたのか、また新型コロナウイルス感染者数とワクチン接種率やその他の要因を結びつけ分析することで、感染者の傾向や今後の推移など考えることが可能ではないかと考えた。

2 データについて

本研究のデータは令和2年2月から令和4年11月までの都道府県ごとの新型コロナウイルス感染者数 [1] や都道府県ランキング [2] などからデータを調査して、研究内容に用いている。各データは以下の通りである。新型コロナウイルス感染者数 (x_1)、ワクチン接種率 (一回目) (x_2)、ワクチン接種率 (二回目) (x_3)、ワクチン接種率 (三回目) (x_4)、喫煙率 (x_5)、平均寿命 (x_6)、生活習慣病による死亡者数 (x_7)、高齢者世帯数 (x_8)、世帯あたりの自動車保有台数 (x_9)、人口密度 (x_{10})、平均気温 (x_{11})、死亡率 (x_{12})、公共交通機関での通勤率 (x_{13})、病院数 (x_{14})。この14要素に加え、新型コロナウイルス第7,8波時の感染者数 (x_{15}) を追加する。

また観測値は人口10万人あたりのものを使用する。分析方法は、相関係数、クラスター分析、重回帰分析を行った。分析に関しては金明哲 [4] を参考に行った。

3 相関係数

まずは新型コロナウイルス感染者数 (x_1) と各データとの相関係数を求めることによって、互いの関係性を調べる。

表1 相関係数

変数	相関係数	変数	相関係数
x_2	-0.7710	x_9	-0.5586
x_3	-0.7668	x_{10}	0.6004
x_4	-0.8485	x_{11}	0.5189
x_5	-0.0087	x_{12}	0.1521
x_6	0.1822	x_{13}	0.6053
x_7	-0.7647	x_{14}	0.5711
x_8	-0.0418		

有意水準5%で棄却された要素は、一回目 (x_2)、二回目 (x_3)、三回目 (x_4)、生活習慣病による死亡者数 (x_7)、世帯あたりの自動車保有台数 (x_9)、人口密度 (x_{10})、平均気温 (x_{11})、公共交通機関での通勤率 (x_{13})、病院数 (x_{14})、であった。

3.1 相関係数に対する考察

新型コロナウイルス感染者数 (x_1) と強い正の相関が見られたのは、人口密度 (x_{10})、平均気温 (x_{11})、公共交通機関での通勤率 (x_{13})、病院数 (x_{14}) があげられる。人口密度 (x_{10}) が高い県では必然的に密になる状況が多いと考えられるため説明ができる。公共交通機関での通勤率 (x_{13}) は、自動車等で通勤している人よりも感染経路が多く、感染者拡大に影響しているといえる。病院数 (x_{14}) は人口が多いところに多く、平均気温 (x_{11}) は沖縄県の感染者が多いのと、東北の県が比較的感染者が少ないことからいえると考ええる。

逆に強い負の相関は、一回目 (x_2)、二回目 (x_3)、三回目 (x_4)、世帯あたりの自動車保有台数 (x_9)、生活習慣病による死亡者数 (x_7) が挙げられる。ワクチン接種率が高い都道府県では感染者数が少ないことから一回目 (x_2)、二回目 (x_3)、三回目 (x_4) は根拠付けられる。

また自動車保有台数が多い都道府県では公共交通機関を用いない通勤や通学が増えるため、このような結果になったと考えられる。しかし生活習慣病による死亡者数 (x_7) が負の相関を示すのは意外だった。がんや糖尿病患者は新型コロナウイルス感染者数 (x_1) との正の相関がみられると思ったが、そこに対しては関係性が感じられなかったため、新型コロナウイルスが体調の悪い人にだけ強く影響しているわけではないと考える。

他のデータから強い相関は見られず、特に喫煙率や高齢者世帯数は多少の負の相関が見られただけで新型コロナウイルス感染者数に対して感染拡大の影響は見られなかった。

4 クラスター分析

次にクラスター分析をウォード法を用いて行った。

4.1 分析結果

第1群: 公共交通機関での通勤率が高い都道府県

第2群: ワクチン接種率が高い都道府県

第3群: 世帯あたりの自動車保有台数が多い都道府県

第4群: 感染者数が最も多い都道府県

第5群: 高齢者世帯数が多い都道府県

4.2 クラスター分析に対する考察

第1群:(愛知(24), 京都(25), 福岡(40) など)

第1群は人口密度(x_{10})と公共交通機関での通勤率(x_{13})が高い都道府県の群である。第4群と似ていて新型コロナウイルス感染者数(x_1)が多く、これは人の移動が多く、人口密度(x_{10})が高いことから説明がつく。また主要都市に近い都道府県や政令指定都市などが多いため電車や地下鉄などの交通機関での通勤率(x_{13})が高いことも理解できる。このことから公共の交通機関での密の状態が感染経路拡大に強く影響していると分かる。

第2群:(青森(1), 岩手(2) など)

第2群はワクチン接種率(x_2), (x_3), (x_4)が他の都道府県と比較し高い群である。東北は高齢者の割合も高く、平均寿命(x_6)が若干低く新型コロナウイルスに対して強く対策する傾向がある。その点同じ東北でも宮城県(5)が別の群なのは東北の中でも栄えているのに理由があると考え、高齢者に対する感染症対策が重要であると改めて理解した。

第3群:(群馬(11), 新潟(16) など)

第3群は世帯あたりの自動車保有台数(x_9)が多い都道府県の群である。都心から離れている群が多く、特に中部地方や関東の北側など地下鉄や路線の少ない都道府県が多く見られた。ただその分、新型コロナウイルス感染者数(x_1)は公共交通機関が多い都道府県に比べ少ないため、密を避ける移動が密を避ける移動が感染経路の縮小につながるといえる。

第4群:(東京(13), 大阪(26), 沖縄(47))

第4群は新型コロナウイルス感染者数(x_1)が多い都道府県の群である。主要都市である東京都(13)と大阪府(27)があり、人口も多く人の移動も多いためこのような結果を得たと考えられる。また沖縄県(47)は2022年の夏に爆発的に感染者が伸びていた。感染経路が多くクラスターの発生も今後も考えらるため意識的な対策をとる必要がある。

第5群:(広島(34), 高知(39), 大分(46) など)

第5群は高齢者世帯数(x_8)が多い都道府県の群である。高齢者の割合が他の群に比べて高く、そのためか第2群と似ていてワクチン接種率が高い数値を示している。ただ三回目(x_4)はまだ十分とは言えない。高齢者の感染は死亡の危険性が高いので接種率を上げていく必要があると考える。

5 重回帰分析

5.1 分析結果

次に、新型コロナウイルス感染者数(x_1)を目的変数、一回目(x_2)から病院数(x_{14})を説明変数として重回帰分析

を行う。変数減少法を行い、最適なモデルとなったものを以下に表に示していく。減少させた結果、三回目(x_4)、平均寿命(x_6)、高齢者世帯数(x_8)、公共交通機関での通勤率(x_{13})、病院数(x_{14})が最適となった。回帰式は以下の通りである。

$$y = -0.757x_4 - 0.173x_6 - 0.196x_8 + 0.219x_{13} + 0.188x_{14}$$

5.2 重回帰分析に対する考察

三回目(x_4)、公共交通機関での通勤率(x_{13})は他の分析から強い影響を与えていると予想できた。平均寿命(x_6)も与えていることから長く生きている人が多い場所でも影響が強く出ていることがわかる。高齢者世帯数(x_8)は感染経路の拡大化や免疫の低い人が多い場所では新型コロナウイルス感染者が増えやすいという結果が得られ、病院数(x_{14})の影響に対してはワクチン接種率だけではない病院の手配や充実化も気を抜けないといえる。

6 まとめ

相関係数、重回帰分析から、都心部での公共交通機関での通勤率(x_{13})や人口密度(x_{10})などの感染経路を直接的に切ることが感染拡大の抑制に繋がっているという結果を得た。またワクチン接種率では三回目(x_4)の影響力が二つの分析を通して強く出ているため、接種率の向上に努めることも必要な要素であるとわかった。新型コロナウイルス第7,8波時の感染者数(x_{15})を目的変数とした場合でも、同様の分析結果が得られた。クラスター分析では各群に属する都道府県に必要な対策を導き出した。都心部では公共交通機関での通勤率(x_{13})が影響を与えて、地方では世帯あたりの自動車保有台数(x_9)などが感染防止に影響しているという結果を得た。

7 おわりに

この研究を行い、新型コロナウイルスに因果関係のある要因をデータと分析を用いて可視化することができた。さらなる変異株の登場も可能性としてはあるため、気を抜かない対策に今までの学びや失敗から取り組むべきである。一刻も早く新型コロナウイルスの恐怖から脱却するため国民一人一人が意識を持った行動を心がけてほしいと願う。

参考文献

- [1] 厚生労働省 オープンデータ (2022/6/24 閲覧)
<https://www.mhlw.go.jp/stf/covid-19/open-data.html>
- [2] 都道府県データランキング (2022/6/26 閲覧)
<https://uub.jp/pdr/>
- [3] COVID-19 対策ダッシュボード (2022/9/18 閲覧)
<https://covid-19.nec-solutioninnovators.com/>
- [4] 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」
森北出版株式会社, 東京, 2007/10/25 出版