

2022年度 統計データ分析コンペティション

審査員奨励賞 [高校生の部]

新型コロナウイルス感染者数減少において
効率的な政策の検討

篠田 桜楽 (品川女子学院高等部)

新型コロナウイルス感染者数減少において効率的な政策の検討

篠田 桜楽
品川女子学院高等部

1 研究のテーマと目的

新型コロナウイルスは2019年12月から感染者が報告されて以来、日本では依然として猛威を振っている。WHOによると感染経路はエアロゾル感染と飛沫感染、接触感染であり、感染が起こりやすい環境条件として「三密」と呼ばれる概念に1つでも当てはまる環境に感染者と感受性者が滞在すると、感染が成立する可能性は高くなり、さらに3つの条件がそろるとより高くなるとある⁽¹⁾。

一方2020年5月に東京都の人口は1400万人を超え、一極集中が進み地方との地域格差が広がっている。また都市圏における都市問題の1つとして公共交通機関の朝の通勤ラッシュがある。JR横須賀線では7:26~8:26の時間帯に混雑率は195%と非常に高い数値をとっている。⁽²⁾そのため日本国内でも新型コロナウイルス疾患の流行し始めた時に感染が懸念される場所の1つとして満員電車が挙げられた。やはり満員電車は「三密」に当てはまるから懸念されたのだろう。都市圏では人が集まりやすいため、「三密」に当てはまる条件可になりやすい。

そこで本研究では新型コロナウイルス感染者数と人口密度の相関について、47都道府県別の新型コロナウイルスの感染者数を目的変数として回帰分析を行い、地域における疾患率の差異と各都道府県が行った政策等から新型コロナウイルス感染の要因分析を試みた。

2 研究の方法と手順

都道府県ごとの新型コロナウイルス感染者数を目的関数、人口を説明変数として単回帰分析を行った後、多数のデータから重回帰分析を行った。なお目的関数には2022年8月5日現在の「累計感染者数」⁽³⁾を用いた。

最初に都道府県ごとの人口と新型コロナウイルス感染者数を単回帰分析することで、新型コロナウイルス感染者数が増加する理由として人口はどの程度相関しているか調べた。その後、多数のデータを用いて、先程調べた単回帰分析の結果と比較して人口以外の社会的事象がどの程度関わっているか調べた。

3 データセットの加工

図1によると、感染者数は都道府県間で差が大きく、上位には東京、大阪、神奈川、愛知、福岡等の人口の多い都道府県が多い。一方で下位には鳥取、山形、秋田と地方の人口が少ない県が並ぶことから、感染者数には人口や「三密」との関係が強いと考えられる。しかし北海道、奈良、沖縄はさほど人口が多くないにもかかわらず感染者数が多いことから、人口関連以外の観光などの要因の関与が示唆される。

説明変数にはSSDSEのデータを主とした総計16種の説明変数を用いた(表1)。なお使用するデータとして「経済論集」⁽⁴⁾を参考に医療・福祉、世帯所得、人口と上記のことから観光等の説明変数を採用した。

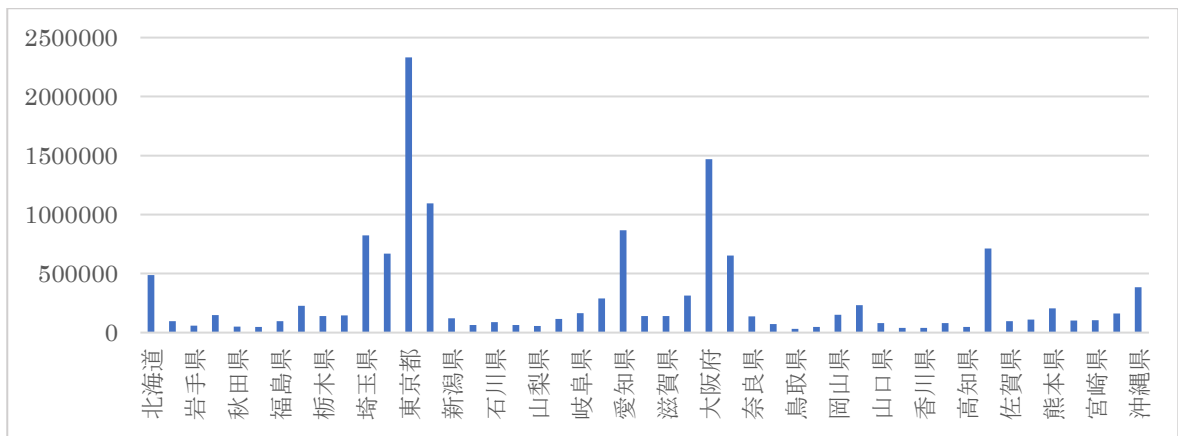


図1 都道府県別累計感染者数(2022/8/5 現在)

表1 解析に用いた説明変数の定義とデータ出典

分野	説明変数名	出典	年度
人口・世帯	総人口, 15歳未満, 15-65歳未満, 65歳以上人口	SSDSE-E-2022v2	2020
	可住地面積	SSDSE-E-2022v2	2020
	消費支出(二人以上の世帯)	SSDSE-E-2022v2	2020
経済	事業所数(民間)(卸売業, 小売業)	SSDSE-E-2022v2	2016
	事業所数(民間)(宿泊業, 飲食サービス業)	SSDSE-E-2022v2	2016
労働	小売店数	SSDSE-E-2022v2	2016
	飲食店数	SSDSE-E-2022v2	2016
	大型小売店数	SSDSE-E-2022v2	2016
娯楽	民間体育施設数	SSDSE-E-2022v2	2018
	劇場, 音楽堂等数	SSDSE-E-2022v2	2018
	一般旅券発行件数	SSDSE-E-2022v2	2020
	延べ宿泊者数	SSDSE-E-2022v2	2020
交通	映画館数	SSDSE-E-2022v2	2016
	公共交通機関通勤・通学率(偏差値)	国勢調査	2015
	乗用車保有台数(偏差値)		2010
医療	一般病院数	SSDSE-E-2022v2	2019

4 データ分析の結果

累計感染者数と人口の単回帰分析を行ったところ下図(図, 表2)の結果となった。

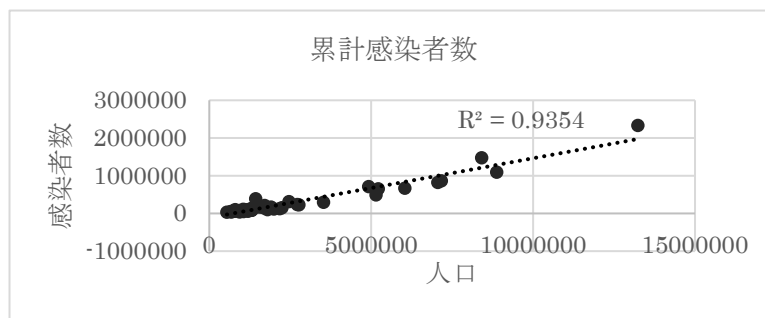


図2 累計感染者数と人口の散布図

表 2 累計感染者数を目的分数，人口を説明変数とした単回帰分析の結果

回帰統計	
重相関 R	0.96981019
重決定 R2	0.9405318
補正 R2	0.93918025
有意 F	1.3327E-28

結果から感染者数と総人口は強い相関があることが分かった。しかし総人口が必ずしも感染者数を決定しているわけではないということも分かった。それは先に述べた通り，人口がさほど多くない地方でも感染者数が多いものもあったからである。そのため次は感染者数を目的関数，様々な分野の 16 項目を説明関数として重回帰分析を行った。結果は下図（表 3）のとおりである。

表 3 感染者数を目的分数，16 項目の説明変数の重回帰分析結果

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	-100651.28	276229.071	-0.3643761	0.71822067	-665603.17	464300.6	-665603.17	464300.6
総人口	0.17850457	0.05634332	3.16815843	0.00359913	0.06326954	0.2937396	0.06326954	0.2937396
可住地面積	-0.0713434	0.10979795	-0.64977	0.52095489	-0.2959054	0.15321861	-0.2959054	0.15321861
公共交通機 関通勤・通 学率（偏差 値）	2976.97623	2976.57088	1.00013618	0.32551721	-3110.7948	9064.74722	-3110.7948	9064.74722
乗用車保有 台数（偏差 値）	1959.64037	2996.24553	0.65403197	0.51824449	-4168.3698	8087.65053	-4168.3698	8087.65053
事業所数 （民営） （卸売業， 小売業）	6.88677348	8.16270777	0.84368737	0.40574991	-9.8078384	23.5813854	-9.8078384	23.5813854
事業所数 （民営） （宿泊業， 飲食サービ ス業）	-35.52875	19.8921227	-1.7860713	0.08454883	-76.212709	5.15520908	-76.212709	5.15520908
民間体育施 設数	-47.937715	96.5902487	-0.4962997	0.62342394	-245.48696	149.611524	-245.48696	149.611524
映画館数	-1453.7642	3997.79966	-0.3636411	0.71876371	-9630.1825	6722.65419	-9630.1825	6722.65419
劇場，音楽 堂等数	613.107553	1799.84642	0.34064437	0.7358268	-3067.9917	4294.20681	-3067.9917	4294.20681
一般旅券発 行件数	0.44406987	1.7085823	0.25990546	0.7967731	-3.0503733	3.93851304	-3.0503733	3.93851304

延べ宿泊者数	0.01154684	0.00517452	2.2314796	0.03354207	0.00096375	0.02212993	0.00096375	0.02212993
小売店数	-34.253739	15.890009	-2.1556778	0.03954208	-66.752457	-1.7550219	-66.752457	-1.7550219
飲食店数	68.1653009	24.4076981	2.79277876	0.00915743	18.2459531	118.084649	18.2459531	118.084649
大型小売店数	-408.3094	342.340836	-1.1926985	0.24265516	-1108.475	291.856225	-1108.475	291.856225
一般病院数	806.000107	244.092116	3.30203253	0.00255348	306.775675	1305.22454	306.775675	1305.22454
消費支出 (二人以上の世帯)	-0.4968656	0.54445051	-0.9126001	0.36897764	-1.6103919	0.61666072	-1.6103919	0.61666072

以上の結果から人口以外に強い相関があるものとして観光業や医療・福祉、飲食店が挙げられる。一方で公共交通機関通勤・通学率や乗用車保有台数、映画館などの娯楽施設数は比較的相関がないと言える。これは長時間滞在やマスクを外す、同じ物を複数人の人で共有するなどの事が起こるからで、公共交通機関利用時や娯楽施設利用時は滞在時間も短く、他人とは違う物の使用や適宜消毒を行っているかだと考える。

よって次に強い相関が出たものの(P値が0.1未満)人口、宿泊業・飲食サービス業の事業所数、宿泊者数、小売店数、飲食店数、一般病院数を説明変数として重回帰分析を行った。結果は下図(表4)である。

表4 感染者数を説明変数、6項目の説明変数を行った重回帰分析の結果

回帰統計	
重相関 R	0.99530174
重決定 R ²	0.99062555
補正 R ²	0.98918332
標準誤差	45387.623
観測数	46

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	-8697.066	19298.7004	-0.4506555	0.65473339	-47732.372	30338.2401	-47732.372	30338.2401
人口	0.12797627	0.01573744	8.13196486	6.3157E-10	0.09614431	0.15980824	0.09614431	0.15980824

事業所数 (宿泊業、 飲食サービ ス業)	-36.965634	15.4695811	-2.3895691	0.02179941	-68.255815	-5.6754525	-68.255815	-5.6754525
延べ宿泊者 数	0.01061991	0.00394606	2.69127011	0.01042892	0.00263825	0.01860157	0.00263825	0.01860157
小売店数	-28.317459	4.57701325	-6.1868859	2.8489E-07	-37.575342	-19.059576	-37.575342	-19.059576
飲食店数	73.1346088	15.349392	4.76465834	2.6242E-05	42.0875329	104.181685	42.0875329	104.181685
一般病院数	934.126306	196.870103	4.74488657	2.7917E-05	535.918937	1332.33368	535.918937	1332.33368

人口、宿泊業・飲食サービス業の事業所数、宿泊者数、小売店数、飲食店数、一般病院数の6項目を説明関数として重回帰分析を行ったことで、16項目で重回帰分析を行った時よりも強い相関が出た。しかし宿泊業・飲食サービス業の事業所数、宿泊者数、小売店数、飲食店数、一般病院数は人口と相関が強くあるため、このような結果がでたと考えられる。そこで人口を目的関数とし、それ以外の5項目をそれぞれ単回帰分析した。結果は下図(図3, 4, 5, 6, 7)のとおりである。

結果から5項目、宿泊業・飲食サービス業の事業所数、宿泊者数、小売店数、飲食店数、一般病院数全て人口と相関があることが分かった。しかし宿泊者数や一般病院数は外れ値が多いことから、人口が少なくはあるが、観光業が盛ん、または高齢者が多く住んでいる地域だと考えられる。また病院数と感染者数の相関がある理由として、多数の病院があることでPCR検査等が行いやすくなるからだろう。そして厚生労働省⁽⁵⁾によると病院や福祉施設では入院患者含め集団感染が起りやすく、感染者数と病院数との相関がある理由として挙げられるだろう。図8によると医療機関や福祉施設の集団感染が上位を占めている。

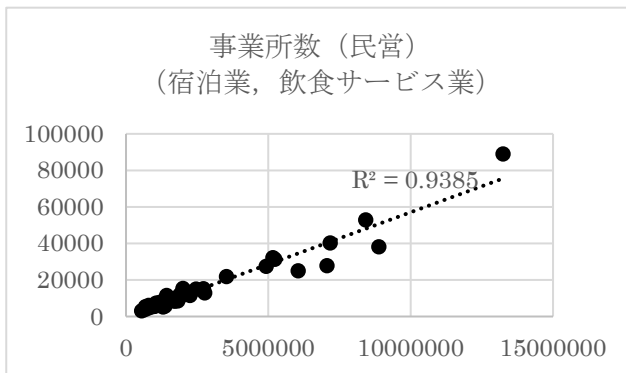


図3 人口と宿泊業・飲食サービス業の事業者数の散布図

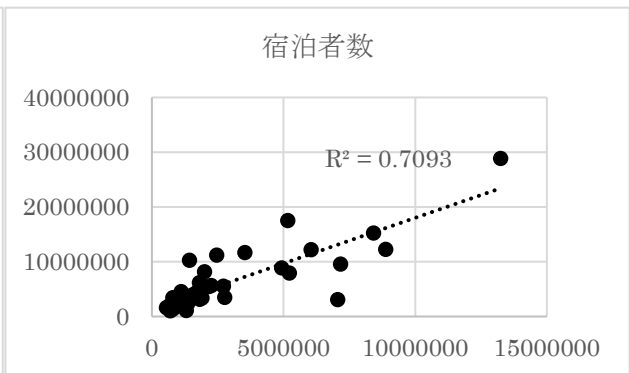


図4 人口と宿泊者数の散布図

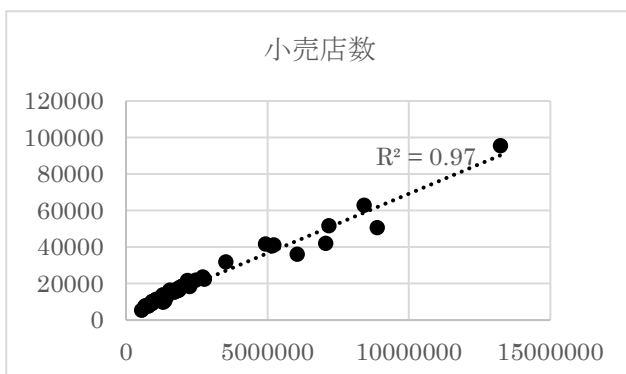


図5 人口と小売店数の散布図

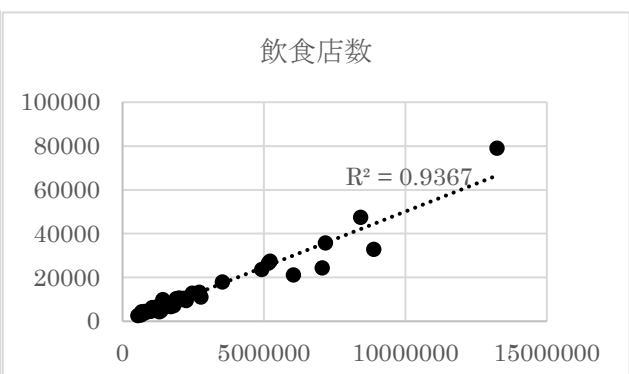


図6 人口と飲食店数の散布図

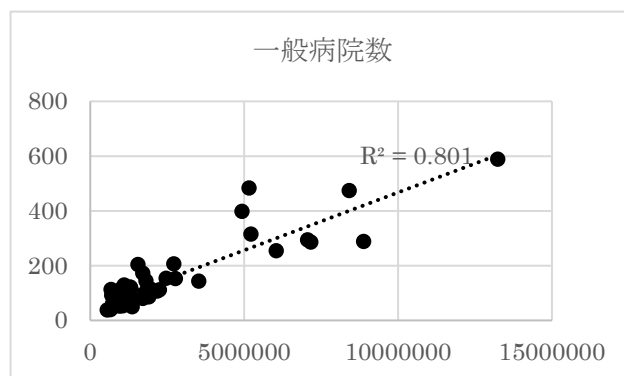


図7 人口と一般病院数の散布図

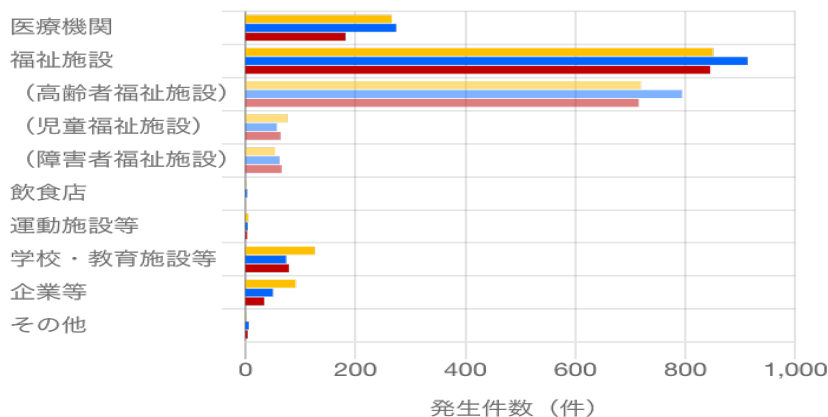


図8 集団感染等発生状況

以上のことから病院数が多いことでPCR検査等が行いやすく、集団感染が起きやすいことが感染者数との相関に繋がり、また病院数が多いことは高齢者数と相関があると考えた。よって感染者数を説明変数、高齢者数を目的変数として単回帰分析をし、次に病院数を説明変数、高齢者を目的変数として単回帰分析を行なった。結果は下図(図9, 10)のとおりである。

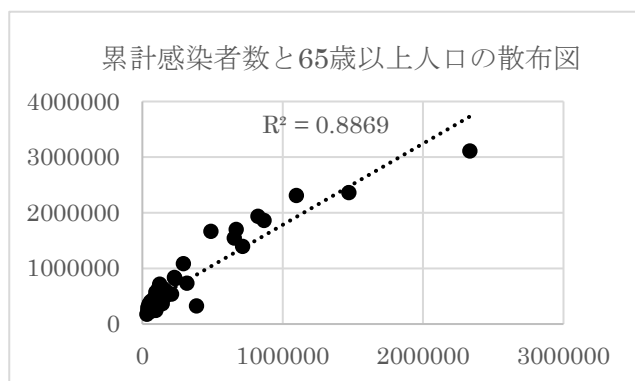


図9 累計感染者数と65歳以上人口の散布図

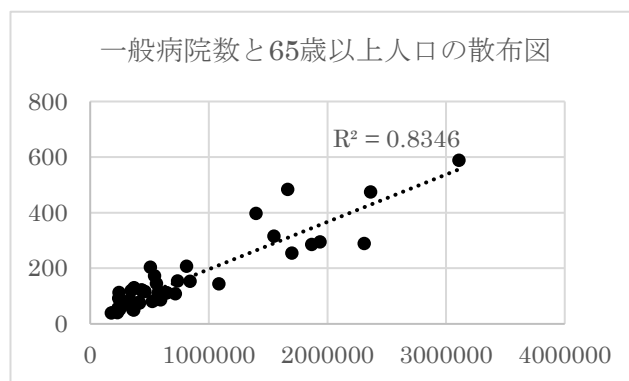


図10 一般病院数と65歳以上人口の散布図

結果から累計感染者数と総人口の相関の方が65歳以上との相関より強いと言える。また病院数と65歳以上人口は強く相関が出ているが、総人口との相関も強いいため先に述べた、病院数が多いことでPCR検査等が行いやすく、集団感染が起きやすいことが感染者数との相関に繋がるという考えは棄却された。では他の年代ごとのデータでは総人口よりも感染者数と相関を疑問に思い、累計感染者数を目的

関数, 15歳未満, 15-65歳未満をそれぞれ説明変数として単回帰分析で行った. 結果は下図(図11, 12)のとおりである.

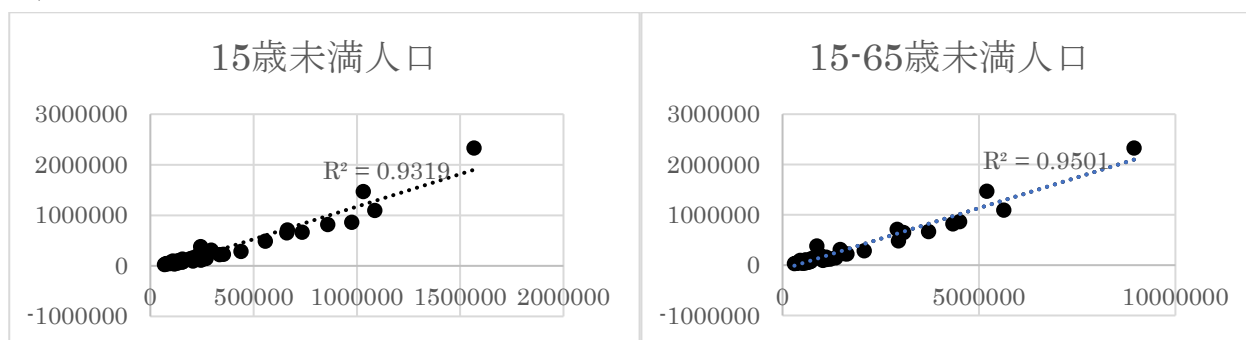


図11 累計感染者数と15歳未満人口の散布図

図12 累計感染者数と15-65歳未満人口の散布図

5 結果の解釈

5-1 考察

世代別人口と累計感染者数との相関は世代ごとの違いがなかったと言える. また新型コロナウイルス感染症対策専門家会議⁽⁶⁾によると

感染を急速に収束の方向に向かわせるためには、人と人との接触を最大限に避けることが必須です。これを、いま集中して実施すべきです。

もし、こうした対策が行われず、人々が何も行動を変化させない場合、感染者数が急増し(赤い上昇線)、一定の潜伏期間後に発症者数も急増することが予想されます(青い上昇線)。(図13)

と説明している。これは結果より娯楽施設よりも長期滞在が見込まれる観光地や医療・介護施設では潜伏期間に感染していると気づかず人と接触することで集団感染が起きやすくなっていると分かっている。症状の軽い人は気がつかないうちに、感染拡大を広げてしまっている。同じく新型コロナウイルス感染症対策専門家会議は「一定条件を満たす場所において、一人の感染者が複数人に感染させた事例が報告されています。」とも説明している。密閉された閉鎖的な空間で多数の人が至近距離で一定時間以上触れ合い、集団感染が発生することで感染の急速な拡大を招くと考えられる。新型コロナウイルス感染は上記の結果から世代関係なく感染するが、比較的症状が軽くなりやすい若者が気づかずに接触し感染者は増加しているようである。

また回帰分析の結果より換気が行われていない密閉された場所では感染率が上がりやすい。以上から新型コロナウイルス対策の効果的な処置としてこれからは個人間の感染防止意識に左右されやすい場所、例えば家族や友人間のホームパーティでの感染防止を呼びかけるべきである。特に基礎疾患があるもしくは関係者にいる場合はより注意を払わなければならない。小売店と感染者数との相関が強かったことから個人間の感染防止対策意識が関わっているだろう。

接触を避けることによる流行拡大抑止効果

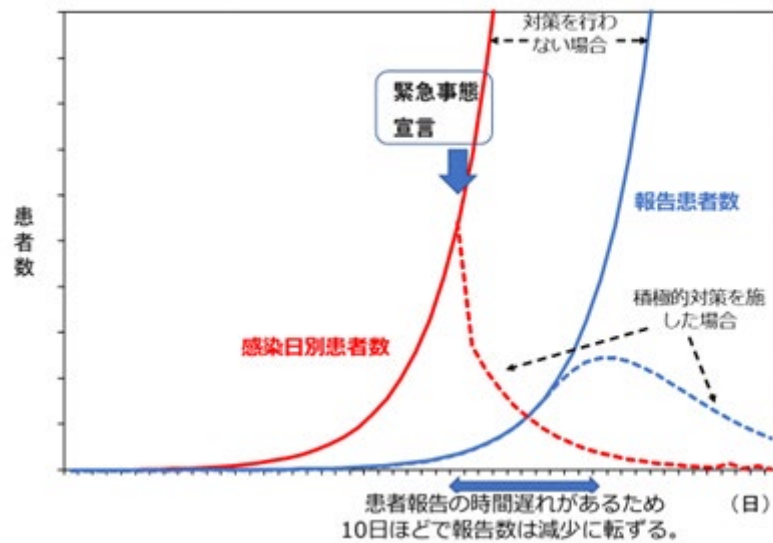


図 13 接触を避けることによる流行拡大抑止効果

5-2 展望

今回行った試行は回帰分析であり、感染者数の相関はわかるものの因果がわかる訳ではない。また重回帰分析を行った結果、人口とは別の社会的指標が感染者数との相関につながっていることが分かる。感染者数と説明変数で重回帰分析を行った説明変数は、全ての項目で人口との相関が見られる。そして各データでの外れ値も検討していき、より精度が高い重回帰分析含めデータ分析を行なっていきたい。

参考文献

- (1) World Health Organization. “Coronavirus disease (COVID-19): How is it transmitted?”
[who.int/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted](https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted)
(参照 2022. 8. 31)
- (2) 都市鉄道政策課. “混雑率データ”. 国土交通省
<https://www.mlit.go.jp/statistics/details/content/001365148.pdf> (参照 2022. 8. 31)
- (3) NHK まとめ. “都道府県ごとの感染者数”. NHK
<https://www3.nhk.or.jp/news/special/coronavirus/data/>
- (4) 鈴木孝弘, 田辺和俊. 国内の新型コロナウイルス COVID-19 感染率の要因分析. 経済論集.
2021年3月. vol. 46. no. 2 (参照 2022. 8. 31)
- (5) データからわかる—新型コロナウイルス感染症情報—. “集団感染者等情報”. 厚生労働省
[Phhttps://covid19.mhlw.go.jp/](https://covid19.mhlw.go.jp/) (参照 2022. 9. 1)
- (6) 新型コロナウイルスについて. “新型コロナウイルス感染症対策専門家会議”. 厚生労働省
https://www.mhlw.go.jp/stf/seisakunitsuite/newpage_00011.html (参照 2022. 9. 1)