

グローバルにみた感染症の家計等の行動への影響：  
機械学習によるアプローチ

国際局 森いづみ、中村俊文\*、乗政喜彦

Bank of Japan Review

2021年5月

新型コロナウイルス感染症の拡大は、グローバルに家計や企業の行動に大きな影響を及ぼしてきた。本稿では、公衆衛生上の措置や人々の移動状況等の大規模データを機械学習によりモデル化し、感染拡大が家計等の行動に及ぼす影響を包括的に捉える「感応度」指標を考案した。この新たな指標からは、①感染症への家計等の感応度は時間経過とともに変化していること、②国・地域間で感応度に大きなばらつきがあること、が示された。また、感染症に対する家計等の行動変化には、非線形性があること——感染が大きく拡大すると、人々の行動は急速に慎重化すること——も示唆された。

## はじめに

2020年以降のグローバルな新型コロナウイルス感染症の拡大は、各国において急速な景気後退をもたらした。昨夏以降、世界経済は、総じてみれば持ち直しから回復に転じているが、そのペースは各国・地域間で大きくばらついている。こうした2020年以降の世界経済の大きな変動や各国間の回復のばらつきには、感染拡大を受けた家計や企業の行動変化が影響していることは言うまでもない。各国で導入された公衆衛生上の措置は家計・企業の行動を制約したほか、感染拡大を受けた人々の自主的な行動変容も、経済に大きな影響を及ぼしたと考えられる。

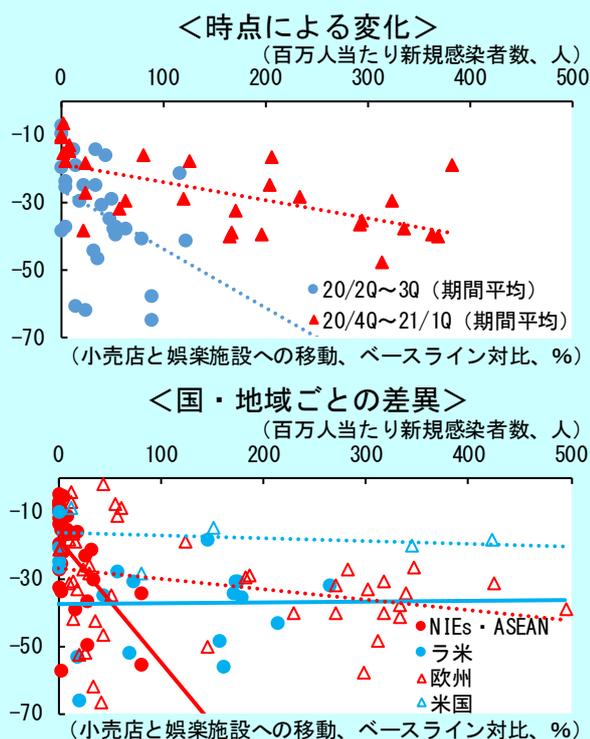
もっとも、こうした感染拡大を受けた家計等の行動変化を包括的に把握することは容易ではない。これは、主として、次の2つの理由による。

1つ目は、この間、家計等の行動は様々な面で変化していることである。感染拡大以降、公衆衛生上の措置や人々の移動状況等を示す高頻度データは急速に整備されたが、これらのうち、どの指標に家計等の行動を表す代表性があるかは明らかではない。そのため、家計等の行動変化を評価する際には、様々なデータを眺めたうえで、総合的に判断せざるを得ないのが現状である。

2つ目は、感染拡大への家計等の反応が、国・地域間でばらつきがあるのみならず、時間を通じても複雑に変化しているとみられることである。例えば、家計の行動を示す代表的な高頻度データである「小売店と娯楽施設への移動状況」と新規感染者数の関係をみると、時期によっても変化しているほか、国毎に大きく異なるといった、データの非線形性が強く窺われる(図表1)。この点、OECDのWeekly GDP Trackerなどの先行研究においても、感染症の与える影響の時系列およびクロスセクションでの極端なばらつきの大きさを指摘している<sup>1)</sup>。このため、単純に複数の指標を合算するなどして、家計等の行動変化を捉えることも難しい。

そこで、本稿では、一つの試みとして、感染拡大を受けた家計等の行動変化について、包括的に把握でき、各国・地域間で比較可能な指標を提案する。具体的には、新規感染者数の増加への反応度合いを捉える「感応度」を、大量の日次データ(125か国の約40種類の変数、約400日分)を用いて算出した。その際には、①どの変数が重要であるか事前には明らかでないこと、②感染拡大への家計等の反応が非線形である可能性があること(例えば、感染がある閾値を超えると、急速に行動が慎重化)、を考慮して、通常の回帰モデルではなく、機械学習の手法を用いたことが大きな特徴である。

【図表 1】新規感染者数と人々の移動状況



- (注 1) PPP ウェイトにおける上位国のうち、30 の国・地域を抽出。NIEs・ASEAN は、香港を除く 7 国・地域ベース。ラ米は、アルゼンチン・ブラジル・コロンビア・メキシコ。欧州は、ベルギー・フランス・ドイツ・イタリア・オランダ・ポーランド・ルーマニア・スペイン・英国。その他インド・ロシア等を含む。
- (注 2) 人口は、2019 年。下段図は、各国・地域の 20/1Q~21/1Q の各四半期平均。20/1Q は 20/2/15~3/31 日。
- (注 3) 図中の線は近似曲線。下段図の赤実線は NIEs・ASEAN、青実線はラ米、赤点線は欧州、青点線は米国。
- (注 4) ベースラインは、20/1/3~2/6 日の該当曜日の中央値。
- (出所) CEIC、HAVER、IMF、UN、Google LLC "Google COVID-19 Community Mobility Reports." <https://www.google.com/covid19/mobility/>. Accessed: 2021/4/26.

## 機械学習を利用した感応度指標

### (定義と考え方)

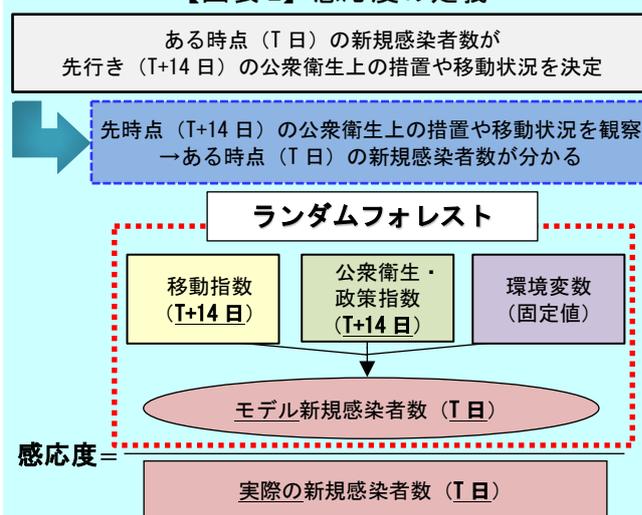
まず、最初に、本稿における「感応度」を定義した後、その意味と解釈について述べる。

本稿で把握することを目指す「感応度」とは、感染拡大（本稿では新規感染者数で捉える）に対する家計等の反応を包括的に捉える指標である。もっとも、前述のように、感染拡大を受けて、家計等の行動は様々な面で変わっているとみられ、単一の変数の変化をもって感応度とするのは適切ではない。そこで、①家計等の行動を示す指標として、どのようなものが考えられるのかを示したうえで、②如何にして感染拡大への諸変数の反応を集約して単一の指標とするか、を検討する。

まず、家計等の行動を示す変数である。この点、国際的なデータの利用可能性も踏まえると、Google が公表している様々な場所への移動指数が候補としてあげられる。また、それに加え、以下の分析では、直接、家計等の行動を示すものではないが、各種の公衆衛生上の措置の状況も対象とした。これは、公衆衛生上の措置が、移動以外の家計等の行動変化を捉える代理指標となりうるためである。

次に、感染拡大のこれら諸変数への影響を如何に集約するか、という点である。諸変数の動きを集約する際には、主成分分析等の線形モデルが用いられることが多いが、本稿の分析には馴染まない。これは、感染拡大と諸変数との関係が先験的に明らかではないもとの、安易に線形関係を前提にモデル化を行うと、変数間の有用な関係を見落とす恐れがあるためである。こうした状況では、事前に変数間の関係性について制約を置かず、最も適する法則をデータに語らせるアプローチとして、機械学習がパフォーマンスを発揮すると考えられる<sup>2</sup>。具体的には、本稿では、ランダムフォレスト（具体的な手順等は BOX を参照）という手法を用いて<sup>3</sup>、以下のような手順を介して、感染拡大の諸変数への影響を集約した感応度を算出した（図表 2）。

【図表 2】感応度の定義



Step 1: ある時点 (T 日) の新規感染者数が、先行き (T+14 日) の公衆衛生上の措置や移動状況に影響を及ぼすと想定<sup>4</sup>。

Step 2: 世界 125 か国のデータを用いて、機械学習（ランダムフォレスト）により、「T 日の新規感染者数」と「T+14 日の様々な公衆衛生上の

措置・移動状況等」との関係モデル化<sup>5</sup>。

—— 機械学習のアルゴリズムに基づき、①「T+14日の様々な公衆衛生上の措置・移動状況等」の諸変数のうち重要なものが採用され、②それらと「T日の新規感染者数」の関係は、非線形性も考慮したうえで、関数推計。

Step 3: それぞれの国について、「T+14日の様々な公衆衛生上の措置・移動状況等」の実績から、モデルが示す T 日の新規感染者数の推計値（モデル新規感染者数）を算出<sup>6</sup>。

Step 4: それぞれの国について、T 日の「実際の新規感染者数」と「モデル新規感染者数」を比較することで、「感応度」を算出。

この「感応度」は、どのように解釈することができるのであろうか<sup>7</sup>。感応度が 1 を上回る場合は、T 日の実際の新規感染者数と比べて、T+14 日時点では、グローバルな平均対比でみて、厳格な公衆衛生上の措置が導入され、移動も減少したことを意味する。つまり、感染拡大に対する感応度が高いと解釈できる。逆に、感応度が 1 を下回る場合は、実際の感染者数と比べて、軽度な措置に止まったり、比較的、移動も自由に行われたことを示す。この意味で、感染拡大に対する感応度は低いと解釈できる。なお、感応度が 1 であることは、当該国が、その時点で感染症に対してグローバルにみて平均的な対応を取っていたことを示している。

### （分析に用いた変数）

続いて、分析に用いた変数について説明する。本稿で用いるのは、125 か国を対象とする、2020 年 2 月中旬から 2021 年 3 月末までの日次ベースのパネルデータである。被説明変数は、新規感染者数であり、特徴量（計量経済学での説明変数に対応）としては、①環境変数（12 種類）、②公衆衛生・政策指数（18 種類）、③移動指数（6 種類）の全 36 種類の変数を利用する<sup>8</sup>（図表 3）。

特徴量の中には、日次で変動しない環境変数もあるが、これらは、モデル上において、ある種のダミー変数として機能していると考えればよい。例えば、「人間開発指数<sup>9</sup>」は、固定値をとるが、先進国ほど高いほか、健康・教育政策を重視する国ほど一層高くなる傾向があるため、先進国と新興国を大別させる役割を担うと考えられる。感応

度の時系列的な動きは、日次で変動しうる「公衆衛生・政策指数」や「移動指数」によって規定されると考えられる。

【図表 3】使用する特徴量（説明変数）

#### （環境変数）

人口密度	年齢の中央値
65歳以上人口割合	70歳以上人口割合
一人当たりGDP	極度の貧困割合
循環器系疾患死亡率	手洗い設備普及率
ベッド数	糖尿病割合
平均寿命	人間開発指数

#### （公衆衛生・政策指数）

学校閉鎖	職場閉鎖
公共イベント中止	集会制限
公共交通機関閉鎖	外出制限
国内移動制限	入国制限
所得支援	家計債務・契約の緩和
財政措置	国際支援
広報キャンペーン	検査方針
接触追跡	医療への緊急支出
マスク着用方針	高齢者保護

#### （移動指数）

小売店と娯楽施設	食料品店と薬局
住居での滞在時間	公共交通機関の駅
公園と屋外	職場

（出所）Our World in Data, Oxford COVID-19 Government Response Tracker, Blavatnik School of Government, University of Oxford.

#### （感応度指標の特徴）

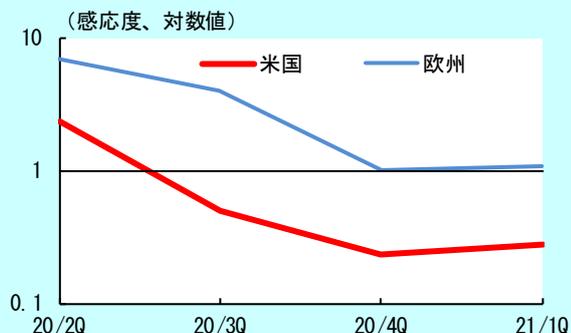
上記の枠組みで推計された感応度について、各国・地域の時系列およびクロスセクションでの特徴について整理する。

まず、感応度の時系列的な特徴としては、全ての国・地域において、感染拡大の当初時点である 2020 年第 2 四半期と比べると、低下している（図表 4）。これは、感染拡大当初は、ウイルスの性質について不確実性が極めて高いもとで、各国において広範かつ厳格な公衆衛生上の措置が導入されたほか、不安心理等から、家計等が自発的に移動を抑制する傾向も強かったことを示していると考えられる。それ以降、昨年末にかけて感応度は徐々に低下したが、これは、公衆衛生上の措置の対象地域や業種を限定する試みが広がるもとで、家計行動の抑制が比較的マイルドとなったことを示しているとみられる。もちろん、こうした家計等の行動変化を受けて、新規感染者数が大きめに増加し、結果として（感染拡大に比べれば、小幅であるにせよ）公衆衛生上の措置が再度強化され、人出が大きめに減少した国も多い。感応度は、あくまで感染拡大への家計等の（短期的な）反応を捉える試みであり、長い目でみた感染症と

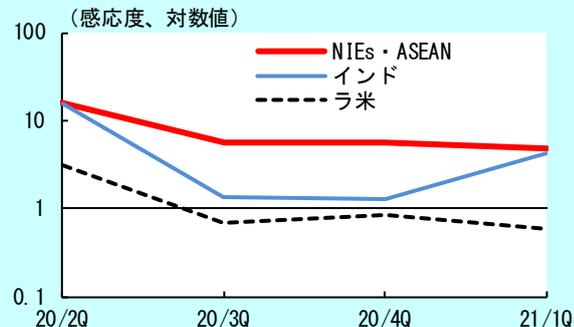
経済との相互作用を捉えるためには、より多角的にデータを分析する必要があるのは言うまでもない。なお、感応度は、本年入り後、幾つかの国・地域で、幾分上昇しているが、これは感染力が高いとされる変異株の拡大への家計等の警戒感の強まりを反映している可能性もある。

【図表 4】 主な国・地域の感応度の推移

(先進国)



(新興国)



(注) 欧州は、フランス・ドイツ・イタリア・スペイン・英国。NIEs・ASEAN は、香港・インドネシアを除く 6 개국・地域ベース。ラ米は、アルゼンチン・ブラジル・メキシコ。地域別は、各国の人口ウェイトを用いて算出。自国を除く 124 か国のデータで学習。scikit-learn により算出。

(出所) Our World in Data, Oxford COVID-19 Government Response Tracker, Blavatnik School of Government, University of Oxford.

感応度の動きからは、各国・地域間での感染症への反応の違いも見取れる。すなわち、先進国の中では、米国と比べて、欧州は、感応度が高めとなっている。また、新興国では、NIEs・ASEAN で感応度が高い一方、ブラジルを中心にラ米では、一貫して低めとなっている。こうした感応度の違いは、NIEs・ASEAN の多くの国が、初期段階から、感染抑制に強めの措置を講じてきたことと整合的である。また、こうした感応度の違いは、新規感染者数が相対的に抑制されている ASEAN の内需が、感染の大幅な拡大に見舞われたラ米と比べても弱いことを、ある程度説明する可能性もあると考えられる。

このように時系列・クロスセクションでの感応

度の動きからは、感染症に対する家計等の反応の変化や各国間の違いを一目で直観的に捉えることが可能である。これは、図表 1 で確認したような散布図等と比べた場合の感応度指標の大きなメリットと考えられる。

## 感染症拡大の影響：どのような行動が、どの程度抑制されたか？

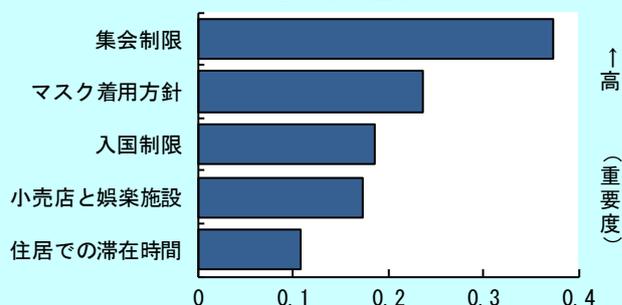
(感染症拡大と行動変化)

次に、前節で推計したモデルを用いて、感染拡大が、どの変数(特徴量)に大きな影響を及ぼしたのか——感染症の拡大は、家計等のどのような行動様式の変容等をもたらしたのか——を考察する。

ランダムフォレストでは、インプットされたどの特徴量を重視して学習したかを、「Permutation Importance (PI: 変数重要度)」という指標を用いて評価する。PI とは、対象とする特徴量を利用しない場合には、モデル精度がどの程度低下するかを表しており、換言すれば、当該「精度の低下分」を特徴量の重要度として解釈して計算された指標である<sup>10</sup>。

PI の高い順に、上位 5 位まで(ダミー変数としての色彩が強い環境変数を除く)を確認すると、①小売店と娯楽施設への移動や、住居での滞在時間といった移動指数に加え、②集会制限やマスク着用方針といった公衆衛生・政策変数も登場している(図表 5)。これは、事前に予想されたように、感染症の拡大は、様々な経路を介して、家計等の行動に影響を及ぼしていることを示唆している。

【図表 5】 変数重要度



(注) Permutation Importance (125 か国のデータで学習) のうち、上位 5 つの特徴量を抜粋(環境変数を除く)。eli5 及び scikit-learn により算出。横軸は、各特徴量の相対的な重要度を表している。

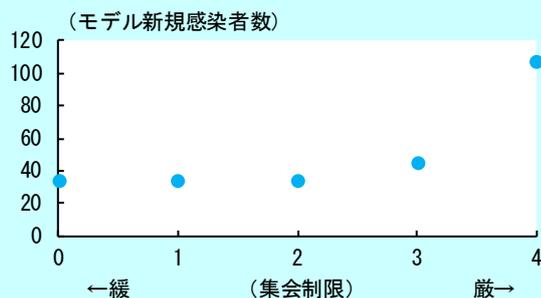
(出所) Our World in Data, Oxford COVID-19 Government Response Tracker, Blavatnik School of Government, University of Oxford.

## （行動様式の変化の非線形性）

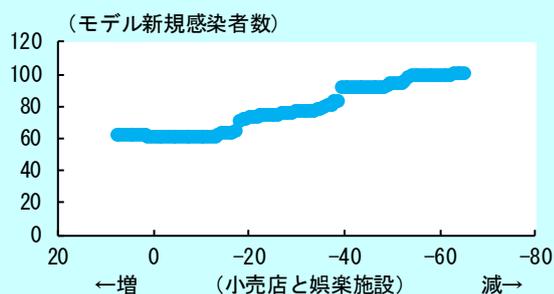
次に、感染拡大による個別の家計等の行動変化の程度を分析する。ランダムフォレストに限らず、機械学習モデルにおける特徴量と結果との関係は非常に複雑であるため、可視化することに困難が伴うが、本稿では Partial Dependence Plot (PDP) を用いて確認していく。PDP とは、ある特徴量と結果の関係について、対象としている特徴量以外をすべて平均化し、ある特徴量の水準毎の結果への影響をプロットしたものである<sup>1)</sup>。PI で上位に登場した特徴量の中から2つを選択して、モデルから示唆される T+14 日の特徴量の水準（家計等の行動、横軸）と T 日の新規感染者数（縦軸）との関係をプロットしたものが、図表 6 である。ただし、ここでの新規感染者数は、他の特徴量を平均化した仮想値であるため、その水準自体に意味がない（実際の新規感染者数を意味するわけではない）点は留意する必要がある。

【図表 6】特徴量の非線形的な影響

（公衆衛生・政策指数：集会制限）



（移動指数：小売店と娯楽施設）



（注 1）PDP（125 か国のデータで学習）をプロット。scikit-learn により算出。縦軸のモデル新規感染者数は、横軸の特徴量以外の特徴量を平均化した場合の値を意味している。

（注 2）集会制限の横軸は、5 段階。0：制限なし、1：制限人数の上限 1,000 人超、2：制限人数の上限が 101~1,000 人、3：制限人数の上限が 11~100 人、4：制限人数の上限が 10 人以下。

（注 3）小売店と娯楽施設への移動の横軸はベースライン対比。ベースラインは、20/1/3~2/6 日の該当曜日の中央値。

（出所）Our World in Data, Oxford COVID-19 Government Response Tracker, Blavatnik School of Government, University of Oxford.

これらの図表からは、感染拡大が、家計等の行動に及ぼす非線形の関係が示唆される。この点についての、一つの（やや大胆な）解釈としては、以下のようなものが考えられる。例えば、集会制限についてみると、モデル上の新規感染者数は、制限が 3 以下の場合、ほぼ横ばいであるが、制限が 4 になると跳ね上がる。このことを縦軸から横軸に解釈すれば、①新規感染者数が一定の水準を下回ってれば、集会制限を強めない国も多いが、②ある水準を上回ると、多くの国で制限が非連続的に高まることを示しているのかもしれない。同様に、小売店と娯楽施設への移動についても、新規感染者数が一定の水準を上回ると、多くの国で非連続的に減少することを示している可能性がある。本稿では、グローバルな平均でみて、「T 日の新規感染者数」と「T+14 日の様々な公衆衛生上の措置・移動状況等」の関係を、ランダムフォレストを利用することで非線形関数も含めて関数近似しており、因果性も含めた予測ではないため、上記のような縦軸から横軸への解釈も許容されうると考えられる。

## おわりに

本稿では、感染拡大への家計等の反応を包括的に捉えることを目的に、機械学習の手法を用いた感応度の指標を新たに考案した。推計値の時系列・クロスセクションの動きをみると、感応度指標は、ある程度、家計等の行動変化を捉えているとみられる。また、この推計に用いた機械学習のモデルからは、感染拡大と家計等の行動変化の間には、非線形の関係があることも示唆された。

以下、本稿の留意点について 3 点述べる。

1 点目は、利用したデータ量についてである。本稿では、日次かつ 125 か国の大量のデータを用いている。とはいえ、全てのデータは「感染拡大後」という単一局面を対象としたものであり、新たにデータが加わった場合、関数全体の推計結果が大きく変化する可能性はある。過学習を防ぐなど一定の対応（回帰木の深さを制限し、多くの回帰木を利用するなど）を取ってはいるが、機械学習による推計結果は、長期的なデータで学習していない場合は汎化が十分ではなく、通常の計量分析と比べて（柔軟に変数間関係を捉えるというメリットがある反面で）データの更新に伴い変化

---

しやすい傾向がある点も否めない。

2 点目は、モデルの妥当性の評価についてである。機械学習による推計の妥当性は、その予測力で評価されることが多い。一方、本稿のモデルは新規感染者数等の予測を目的としたものでないため、その妥当性について統計的な評価を行っていない。こうした点からは、算出した感応度指数の経済学的な意味合いについての評価など、予測精度以外のモデルの有用性の検証手法の在り方について発展させる必要がある。

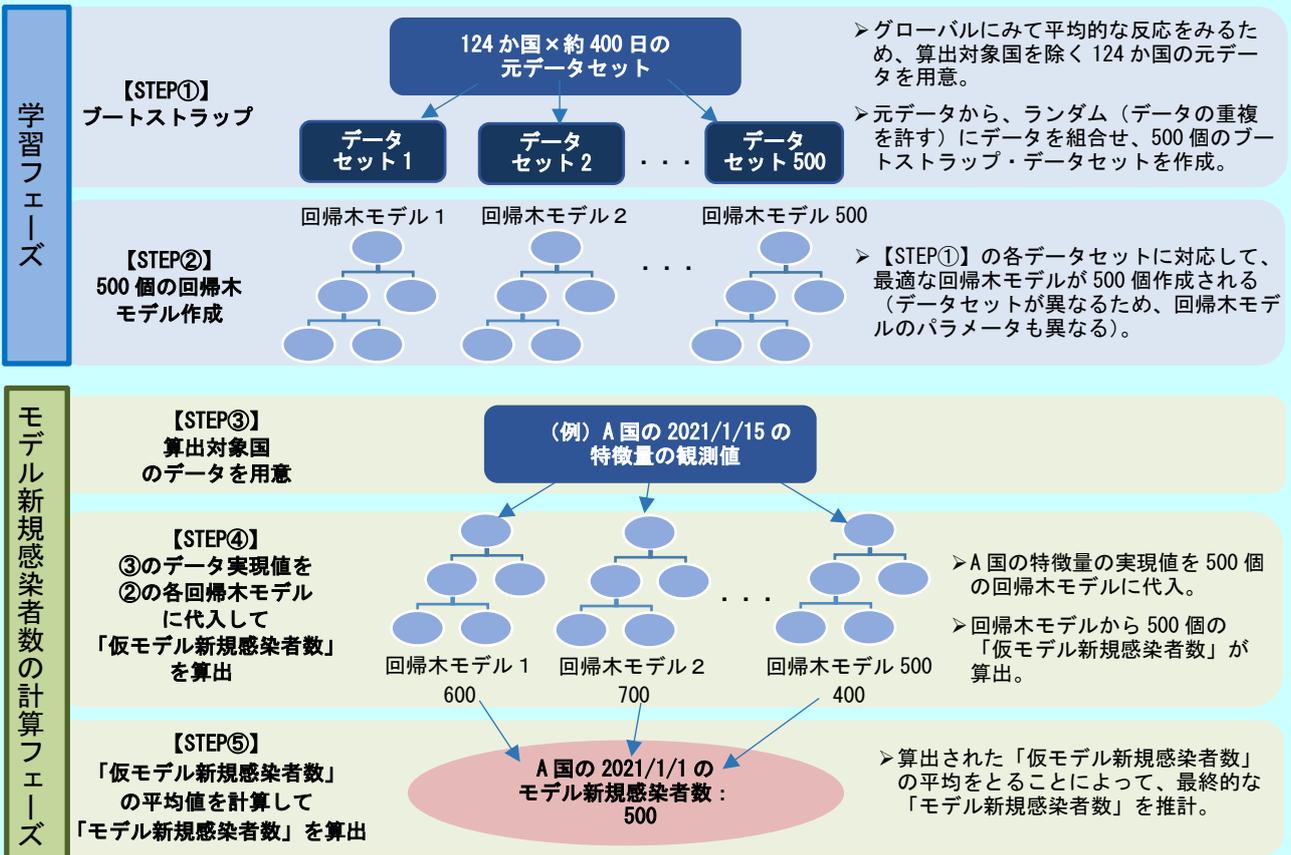
3 点目は、その他（ニューラル・ネットワーク等）の機械学習手法を用いた比較検証の余地が残されている点である。また、感染拡大の家計行動等への非線形的な影響に関する解釈の頑健性についても、複数の可視化手法を試すなどの拡張が

考えられる。最近では、機械学習的なアプローチと計量経済学的なアプローチを組合せ、統計的な推定量を求めることが可能なアプローチも提唱されており、こうした展開も追いながら、分析手法の改善を進めていく必要もある<sup>12</sup>。

冒頭で指摘したように、本稿で提示した感応度の指標は、現時点では、一つの試みであり、改良点も多いと思われる。今後とも、こうした留意点も踏まえながら、モデルの特性や、結果の解釈を丁寧に点検していきたい。また、感染拡大の家計等の消費行動などの経済活動に及ぼす影響を判断していくためには、当然ながら、本稿で試算した感応度指標だけではなく、本稿で推計に用いなかった指標も含め、多くの指標を見ていくことも重要である。

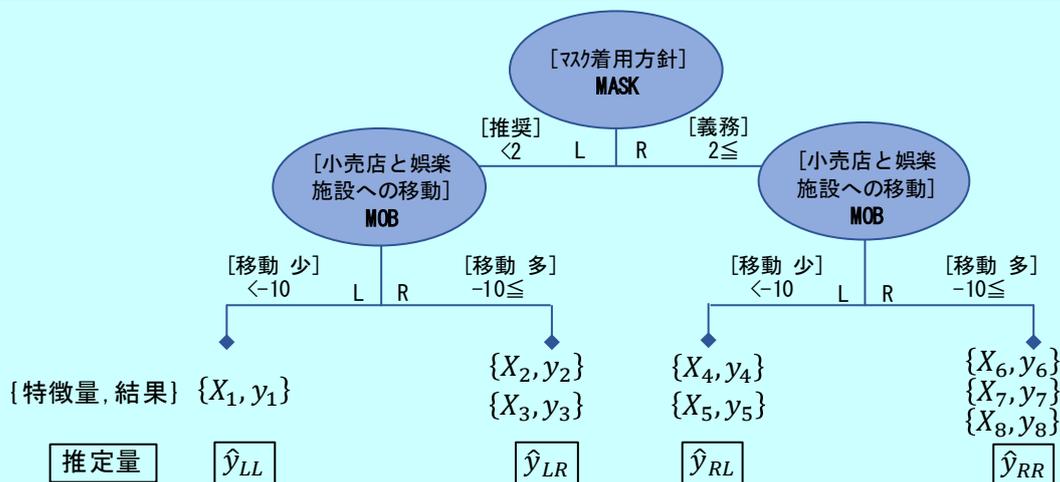
1. 学習フェーズ：世界 124 か国（除く算出対象国）の新規感染者数と特徴量（約 400 日分）を元データとして、ブートストラップ法で 500 個の仮想的なデータセットを作成【STEP①】。各データセットを学習して、500 個の回帰木が作成される【STEP②】。
2. 計算フェーズ：算出対象国の観測値（ある時点）をインプットすると【STEP③】、各回帰木から 500 個の「仮モデル新規感染者数」が算出され【STEP④】、当該 500 個の平均値が「モデル新規感染者数」としての推計値となる【STEP⑤】。

＜ランダムフォレストによる学習・推計のイメージ＞



＜（参考）回帰木のイメージ：深さ 2 のケース＞

CART (Classification And Regression Tree) アルゴリズムに基づき、①特徴量（例：MASK）の種類 ( $X_i$ )、②分岐に登場する特徴量の順番、③分岐の際に用いる閾値といったパラメータが決定され、分類された特徴量と結果（新規感染者数： $y_i$ ）との関係が回帰分析により推定される。



\*現・政策委員会室

<sup>1</sup> グローバルな感染拡大と経済活動の関係を分析した文献として、例えば、Woloszko(2020)が挙げられる。

Woloszko, N. (2020), "Tracking Activity in Real Time with Google Trends", OECD Economics Department Working Papers No.1634.

<sup>2</sup> 中央銀行におけるビッグデータおよび機械学習の利用状況や課題認識については、例えば、Doerr, et al. (2021)が参考になる。

Doerr, S., Gambacorta, L., and Serena, J.M. (2021), "Big Data and Machine Learning in Central Banking", BIS Working Papers No.930.

また、日本銀行においても、以下の日銀レビューやワーキングペーパーで機械学習の可能性について考察している。

三木・五島、「機械学習と株価予測—ニューラル・ネットワークの活用—」、2020年11月、日銀レビュー。

驚見、「通貨オプション市場における投資家センチメントの要因分析：機械学習アプローチ」、2020年10月、日本銀行ワーキングペーパー・シリーズ。

<sup>3</sup> 本稿では、数ある教師あり機械学習モデルの中から、以下の2点を考慮し、ランダムフォレストを採用した。第1は、非線形性が問題となり得るデータ分析に有用な点である。感染拡大によって、家計等の行動が非連続的に変化することを踏まえると、この点を容易に捉えやすいことは大きなメリットである。第2は、過学習を抑制するためのパラメータ調整が容易な点である。本稿の目的は、グローバルにみた平均的な反応と当該国の反応を比較して、感応度を算出することであり、グローバルにみた平均的な反応は、当該国に特有の傾向を過度に反映しないことが望ましい。

なお、機械学習の中でのランダムフォレストの位置付けについては、Mullainathan and Spiess (2017)では、局所的ノンパラメトリックな関数クラスに属し、過学習を抑制する正則化パラメータ (regularizer) として「木の分岐の深さ (depth)」と「木の数 (bootstrap の回数)」をもつアルゴリズムとして整理している。また、ランダムフォレストの弱点として、Varian (2014)は、ブラックボックスの要素が残る点を指摘している。すなわち、一本の決定木であれば、特徴量の役割を一目で理解できるが、複数の決定木からの平均値となると解釈が困難になるためである。

Mullainathan, S., and Spiess, J. (2017), "Machine Learning: An Applied Econometric Approach", *Journal of Economic Perspectives*, 31.

Varian, H. R. (2014), "Big Data: New Tricks for Econometrics", *Journal of Economic Perspectives*, 28.

<sup>4</sup> 新規感染者数の変化が家計行動等に及ぼすタイムラグについては、幾つかのケースを想定して機械学習で推計したうえで、テストデータにおいて最も説明力の高い14日間とした。

<sup>5</sup> ランダムフォレストの計算は、scikit-learn のアルゴリズムを利用した。

Regregosa et al. (2011), "Scikit-Learn: Machine Learning in Python", *Journal of Machine Learning Research*, 12.

<sup>6</sup> 正確には、ある国のモデル新規感染者数の算出には、「算出対象国を除く124か国」のデータで学習し、新規感染者数と公衆衛生上の措置・移動状況等との関係をモデル化している。

<sup>7</sup> T日の新規感染者数は、先行きT+14日の公衆衛生上の措置・移動状況等に影響するが、①両変数の間に逆の因果性はなく、②先行きの期待を織り込んでT日の新規感染者数が変化しない、との前提のもとで、「感応度」を解釈している。

<sup>8</sup> 本稿では、新規感染者数、環境変数、移動指数は Our World in Data の website、公衆衛生・政策指数は Oxford COVID-19 Government Response Tracker, Blavatnik School of Government, University of Oxford から取得できるデータを用いて分析を行った。

<sup>9</sup> 人間開発指数 (Human Development Index: HDI) は、国連が発表している、「出生時平均余命」、「就学予測年数・平均就学年数」、「一人当たり GNI」を指数化したもので、ある国の豊かさに加え、教育や保健医療分野への政策の運用が加味された指標である。

<sup>10</sup> Permutation Importance (PI) は、機械学習のアルゴリズムに依存せずに、特徴量が予測値に与える重要性を統一的に評価できる代表的な指標である。PI の考え方はシンプルであり、特徴量をランダム化することで増加した誤差分が、当該特徴量が精度向上に貢献したと解釈し、重要度と捉えるものである。なお、相関の高い特徴量を同時に機械学習モデルで学習させる場合には、何れかの特徴量の重要度が低下するといった問題が発生する点には留意する必要がある。なお、本稿では、SHapley Additive exPlanations (SHAP) 等の複数の手法で特徴量の重要度を計測したが、PI の結果と概ね同様の結果が得られている。

<sup>11</sup> Partial Dependence Plot (PDP) は、機械学習モデルのブラックボックス化を解消するために Friedman (2001) が提案した手法であり、特徴量と予測値との平均的な関係を可視化する代表的な手法の一つとして利用されている。機械学習モデルのブラックボックス化に対して、特徴量と予測値との関係を可視化する方向での取り組みについては、Goldstein et al. (2015) を参照。

Friedman, J.H. (2001), "Greedy Function Approximation: A Gradient Boosting Machine", *The Annals of Statistics*, 29.

Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E. (2015), "Peeking inside the Black Box: Visualizing Statistical Learning with Plots of Individual Conditional Expectation", *Journal of Computational and Graphical Statistics*, 24.

<sup>12</sup> 例えば、Athey et al. (2019) により提唱された Generalized Random Forests では、特徴量 (説明変数) から予測値に対する効果について、GMM 推定の枠組みにおいて、ランダムフォレストを適用し、推定量および推定量の漸近的な性質を明らかにしている。また、機械学習アプローチ全般における推定量の問題点については、Leeb and Pötscher (2006) が参考になる。

Athey, S., Tibshirani, J., and Wager, S. (2019), "Generalized Random Forests", *The Annals of Statistics*, 47.

Leeb, H. and Pötscher, B. M. (2006), "Can One Estimate the Conditional Distribution of Post-Model-Selection Estimators?", *The Annals of Statistics*, 34.

<sup>13</sup> 本稿では、正則化手法である①フォレストの中の回帰木の数 (ブートストラップ数) を 500 回、②各回帰木の分岐の深さ (depth) を 10 に設定している。なお、Biau and Scornet (2016) は、①については、計算負荷の制約の中で出来るだけ大きくすることを推奨する一方、②について、先験的には決められないとしている。同論文では、ランダムフォレストにおける推定量の性質についても詳細に検討している。

Biau, G. and Scornet, E. (2016), "A Random Forest Guided Tour", *TEST*, 25.

日銀レビュー・シリーズは、最近の金融経済の話題を、金融経済に関心を有する幅広い読者層を対象として、平易かつ簡潔に解説するために、日本銀行が編集・発行しているものです。ただし、レポートで示された意見は執筆者に属し、必ずしも日本銀行の見解を示すものではありません。

内容に関するご質問等に関しましては、日本銀行国際局国際調査課 (代表 03-3279-1111) までお知らせ下さい。なお、日銀レビュー・シリーズおよび日本銀行ワーキングペーパー・シリーズは、<https://www.boj.or.jp> で入手できます。