

特集：ビッグデータと保健医療の未来

データサイエンスの保健医療への応用

大西 立顕*

抄 録

過疎化や働き手不足により、日本全国で買い物や医療へのアクセスに困難を抱える人が増大している。データサイエンスの保健医療への応用として、本研究では国勢調査と座標付き法人電話帳の100mメッシュデータを用いて、住民の居住地と店舗・施設の位置情報を組み合わせることで、全国の国民一人一人について最寄り店舗・施設までの距離を推定した。2010年と2015年の2時点を比較した結果、5年間で全年代では最寄り店舗・施設までの距離が縮小している一方で、高齢者では相対的に距離が遠くなっていることを明らかにした。推計した距離を都道府県や市区町村単位で年齢階層別に集計し、距離の比や変化率を用いることで買い物や医療へのアクセスが困難である地域を検知した。さらに、高齢者の生活環境を評価し、悪化や改善の傾向を明らかにした。

キーワード：地理空間ビッグデータ、電話帳データ、買い物困難、医療困難、最寄り店舗・施設

社会保障研究 2024, vol.9, no.1, pp.33-44.

I データサイエンスの潮流

コンピューターの性能はムーアの法則に従って向上しており、18カ月ごとに二倍になるというペースで指数関数的に進化している¹⁾。これは10年で100倍、20年で1万倍、30年で100万倍という驚異的な性能向上になる。実際、20年前のコンピューターの性能と比較すると、計算時間やデータ容量の面でも著しい進歩が見られる。例えば、20年前には1年かかるとされる計算が、いまでは

わずか1時間で可能になっている。パソコンの主記憶装置も2000年頃の数ギガバイトのハードディスクから、現在では数テラバイト²⁾のSSDが主流になっている。情報通信ネットワークも2000年頃の3Gと呼ばれる毎秒数メガビットの通信速度から、現在では毎秒数10ギガビット³⁾の5Gに向上している。このような指数関数的な性能向上は物理世界では考えられない。例えば、自動車や新幹線が20年後に1万倍のスピードで走るようなことは起こり得ない。物質世界では物理法則や物理的制約が進歩のペースを制限するが、情報の世界はそ

* 立教大学大学院人工知能科学研究科 教授

¹⁾ 厳密にはこれは半導体素子に集積されるトランジスタ数についての話であるが、コンピュータの処理能力の向上は、主に半導体の集積技術の向上に依存している。

²⁾ 1テラバイトは1000ギガバイトである。

³⁾ 1ギガビットは1000メガビットである。

のような制約を受けにくいことを、コンピューターの驚異的な性能向上が実現している。2000年頃にはコンピューターは多くの業務で実用的な性能を持つようになり、その後の進展は生活に大きな変革をもたらし、現代の生活を根底から変えるものになってきている。

このように情報通信技術が進展したことにより、社会のあらゆる領域で日々、多様なデータが生み出され、その量は爆発的に増加し続けている。特に人文社会科学系分野における研究では、従来はアンケート調査や実験などを通じて目的を持ってデータを収集し、そのデータを理論や仮説に基づいて検証することで分析が行われてきた。しかし、現在では、(多くの場合、異なる目的で)自動的に収集されたビッグデータを活用することで、これまでに調査できなかった対象に対しても分析を行うことが可能になっている。既存の理論や仮説にとらわれることなく、データドリブン(データ駆動)で分析を行うことも可能になり、これまでにない新たな発見が生まれやすくなっている。さらに、計算機の性能が向上したことにより、線形回帰などに代表される単純な分析だけではなく、非線形性、非正常性、非正規性などデータが持つ複雑で本質的な性質を反映させた分析も容易に行えるようになってきている。

経済物理学の分野では、1990年代から大規模なデータに基づいた実証的な分析が積極的に行われている。当初は、すでに大量のデータが利用可能であった株式や為替などの金融市場が主な分析対象であったが、その後、個人や法人の所得、企業間の取引ネットワーク〔大西(2021), p.66〕〔藤原(2024), p.4〕、不動産取引〔大西・渡辺(2019), p.15〕など、さまざまなデータの分析が行われている。ブログやTwitter(X)、TVメタデータ〔Ito and Ohnishi(2021), p.1〕などのテキストデータから社会を分析するような研究も行われるようになり、社会物理学や計算社会科学といった新たな分野も誕生している。地理空間ビッグデータを活用した分析も進展している。一般に、地理空間情報と呼ばれる空間データを扱うための情報量や計算コストは、時系列データやネットワークデータ

(多くの場合、疎行列で表現される)よりもはるかに大きくなる。そのため、時空間ビッグデータの観測や分析はこれまでは容易ではなかったが、計算機の能力向上により実行しやすくなってきている。例えば、スマートフォンや携帯電話の位置情報を活用して全国規模で流動人口を把握し、昼間人口と夜間人口の差から各地域の住宅地の場所を特定した上で各住宅地から外出した人の人数を推定するようなことも可能であり、このような観測に基づいて新型コロナウイルス流行下での地域住民の自粛率を可視化する研究も行われている〔Mizuno et al.(2021), p.453〕。また、電話帳データを利用して数ヶ月単位で市区町村やより小さな地域ごとに店舗数の増減を計測することで、新型コロナウイルスに関連した倒産が全国各地でどの程度起こっているかを推定するような研究も行われている〔Saito et al.(2022), p.545〕。

本稿では地理空間ビッグデータを活用した具体的事例として、食料品や医療施設へのアクセスが困難な地域を推定した分析を紹介する。元々は電話番号を調べるために作成された電話帳のビッグデータを活用することで、買い物が困難な地域を推定するといった異なる目的にデータを有効活用できることを示す。一方で、データを活用する際には多くの場合、サンプルバイアスの問題が生じる。これに関連した分析の限界についても触れる。

II 地理空間ビッグデータを用いた買い物困難・医療困難地域の推定

1 研究の背景

近年、過疎化や働き手不足により、店舗や公共機関の減少、公共交通の路線縮小や本数削減が相次いでいる。これにより、人口減少や超高齢化が深刻な地域だけでなく都市部においても、買い物や医療に困難を抱える人が増加している。生鮮食料品を販売する店舗へのアクセスが困難な状況にある人のことを買い物困難者と呼ぶ。日本全国の買い物困難者は、経済産業省によると約700万人〔経済産業省(2015)〕、農林水産政策研究所による

と約824.6万人に上ると推定されている〔農林水産政策研究所（2018）〕。この問題は「買い物難民」、「買い物弱者」、「食料品アクセス困難」、「フードデザート」といった用語でも表現され、住民の生活環境の維持や高齢者の介護予防にも直結する社会問題として認識されつつあり、近年、この問題に関するさまざまな研究が行われている。

近隣のスーパーマーケットが撤退したことで自家用車を持たない高齢者が満足に買い物に行けなくなってしまったという話を聞いたことがある人は多いであろう。この事例に象徴されるように、問題の本質的な要因は社会的弱者が集まって暮らす地域において生鮮食料品の供給環境が悪化することであると考えられる〔岩間（2013）〕。地方や農村、都市圏といった地域の特性によって問題の背景や状況は異なっており、家族や地域社会の助け合い（ソーシャル・キャピタル）の低下が問題をさらに深刻化させていることも指摘されている〔岩間（2017）〕。この問題に対処するためには、どの地域でどの程度の問題が生じているかの実体を把握することが重要であり、まずは、自宅から最寄りの生鮮食料品店へのアクセスの困難さを指標として定量化することが必要になる。全国規模のこのような指標として、買い物困難者の人口を日本全国の500mメッシュの単位で推定して可視化した「食料品アクセスマップ」が知られている〔農林水産政策研究所（2018）〕。これは、店舗までの距離が500m以上でかつ自動車の利用が困難な65歳以上の高齢者を買い物困難者として定義し、2015年国勢調査と2014年商業統計の地域メッシュ統計（500mメッシュ統計）からデータを収集して推定されたものである〔薬師寺（2015）〕。一方、特定地域のみを対象とした分析についても山形県〔三浦・古藤（2010）、p.643〕や福岡県福岡市〔鳥海（2014）、p.993〕、山形県鶴岡市〔中村・渡辺（2014）、p.33〕、東京都港区〔中村・浅見（2019）、p.437〕、富山県富山市〔衣笠他（2023）、p.735〕など、多くの研究が行われている。これらの研究では、単に店舗へのアクセスの困難さだけではなく、アンケート調査を加えたり、食生活と栄養面から店舗の食料品の品揃えも考慮するなど〔岩間

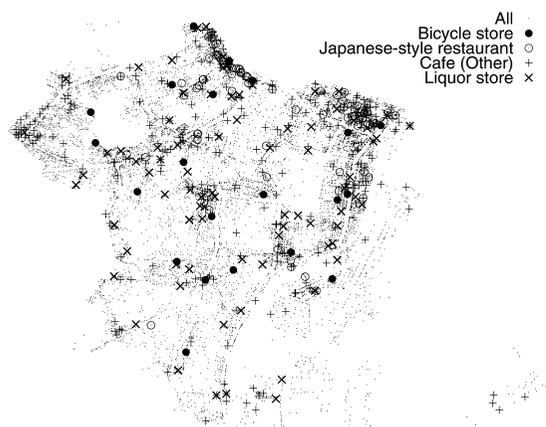
他（2018）、p.81〕、対象は一部の特定地域に限定されているがより実態を反映させた分析も行われている。

日本全国の規模で買物が困難である地域を特定するには、理想的には全国規模でアンケート調査を実施する必要があるが、現実には膨大なコストがかかるためそれを行うのは難しい。実際、全国規模で買い物困難地域の地理的分布を推定した調査は、農林水産政策研究所による食料品アクセスマップのみが知られている〔農林水産政策研究所（2018）〕。このアクセスマップでは、メッシュデータから推定を行っている都合上、居住地から500m以内に店舗があるか否かの情報のみをもとに買い物困難者の人口を推計している。閾値を500mとして固定して調べているため、100mの距離の人と400mの距離の人、あるいは800mの距離の人と2kmの距離の人などの差異を考慮した考察を行うことができない。実際、適切な閾値は、地域の特性や対象とする人の属性によって大きく異なってくると考えられる。さらに、国勢調査と商業統計は数年に一度しか調査が実施されないため、両者の調査が同時点に行われるとは限らずデータの時点にずれが生じてしまう問題、最新のデータが利用できない問題もある。

本研究では、国勢調査100mメッシュ推計データと座標付き電話帳DBテレポイント法人版のデータを活用することで、100mという高い精度で最寄り店舗までの距離を推計する。これにより、全国の住民一人一人に対して店舗・施設までの距離をより詳細に分析することが可能になる。電話帳データは数ヶ月ごとに更新されるため、国勢調査が実施された時点と同時点の店舗・施設の情報を得ることができる。さらに、最新の店舗・施設の所在地データを用いた分析も可能であり、食料品店のみに限ることなく商業統計の調査対象になっていない業種に対しても同様の分析を行うこともできる。

2 国勢調査100mメッシュ推計データと座標付き電話帳DB

本研究では、2010年と2015年の国勢調査100m



注：●は自転車店，○は酒店，+は喫茶店（その他），×は料亭を示す。

図1 電話帳データに収録されている東京都港区の店舗・施設の空間分布

メッシュ推計データを使用し、全国規模で住民の居住地を把握した。このデータは、地域メッシュ統計の4次メッシュ（500mメッシュ）集計データの人口を、住居系建物の戸数に基づいて100mメッシュに配分することで得られるものであり、各100mメッシュ単位で0～14歳人口、15～64歳人口、65歳以上人口、75歳以上人口、85歳以上人口、高齢単身世帯数、高齢夫婦世帯数が推計されている。全国において人口が1人以上いるメッシュの数は、2010年時点で3,844,802個になる。

食料品・飲食店や病院・医療施設の所在地を特定するためには、座標付き電話帳DBテレポイント法人版（2010年2月、2015年2月）を使用した。このデータは、電話帳に収録されている店舗・施設に経度と緯度の座標値、そして業種の情報が付加されたものである。図1は、東京都港区における電話帳に収録されている全店舗・施設の位置を点として示したものである。電話帳には、大企業から個人事業主まで幅広い店舗・施設が収録され

ているため、全国規模であらゆる業種の店舗・施設の所在地を特定することが可能である。業種は39の大分類、824の中分類、2209の小分類に分類されており、それぞれの店舗・施設に記載されている。本研究では、食料品・飲食店に該当する71個の中分類業種と、病院・医療施設に該当する32個の中分類業種を抽出して分析を行った。その結果、2010年時点で抽出された食料品・飲食店は1,078,381個、病院・医療施設は356,267個であった。

3 最寄り店舗・施設までの距離

全国の国民全員に対して最寄りの食料品・飲食店までの距離を調べた。人々の居住地は100mメッシュデータで提供されているが、食料品・飲食店の位置情報はポイントデータ（経度と緯度）であるので、両者のデータ形式を一致させる必要がある。本研究では、居住地の経度と緯度を各メッシュの中心点として定義することで、メッシュデータをポイントデータとして扱った。各居住地（つまり、各メッシュの中心点）から最寄りの食料品・飲食店までの距離を計算するためには、その居住地から最も近い店舗を特定する必要がある。そのためには、各居住地からすべての食料品・飲食店までの距離を計算し、最小距離の店舗を見つける必要がある⁴⁾。この計算には膨大な計算量が必要になるが、並列計算に適した計算になるため、MPIを用いてスーパーコンピュータで実行した。また、本研究では全国規模で分析を行うため、経度と緯度から二地点間の距離を単純に計算すると、緯度によって実際の距離が異なってしまう問題がある。このため、ヒュベニの公式⁵⁾を用いることで、軽い計算コストで比較的高い精度で距離を算出した。

群馬県北部の一部の市町村において、各居住地から最寄りの食料品・飲食店までの距離を計算し

⁴⁾ 実際には計算効率を向上させるために、アルゴリズムを工夫して実行している。具体的には、計算を開始する前に、すべての居住地と食料品・飲食店を経度の昇順に並べ替える。次に、注目している居住地の最寄りの食料品・飲食店を見つけるために、その居住地の経度に近い順に食料品・飲食店を優先的に選択し、距離を計算していく。この際、ある一定の距離以上離れた食料品・飲食店は、順番の情報だけからその距離がすでに計算された距離よりも大きくなるのが分かるため、計算コストのかかる距離の計算を省略することができる。

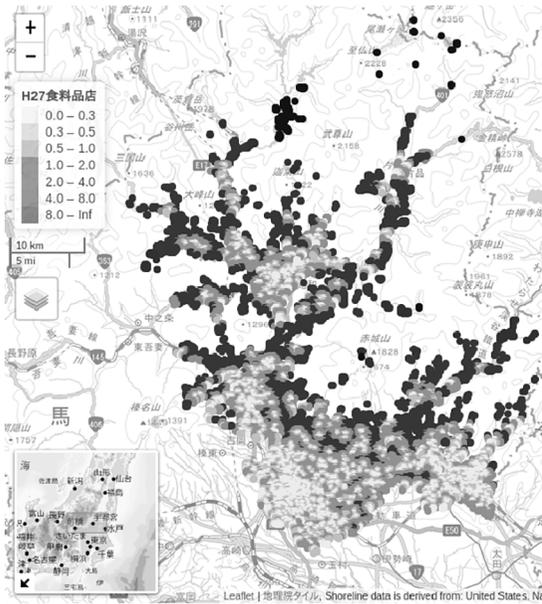


図2 群馬県北部の一部の市町村について居住地から最寄りの食料品・飲食店までの距離 (km) を示したマップ

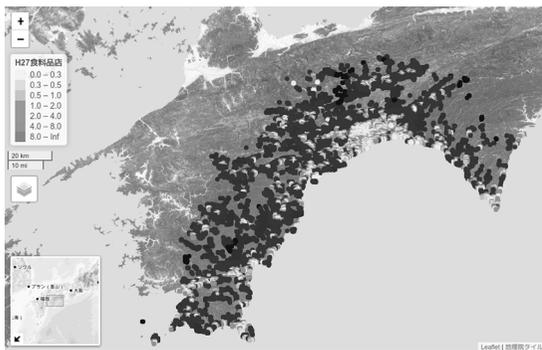


図3 高知県について居住地から最寄りの食料品・飲食店までの距離 (km) を示したマップ

た。図2にその結果を示す。都市部では、最寄りの食料品・飲食店までの距離が300m以下の居住

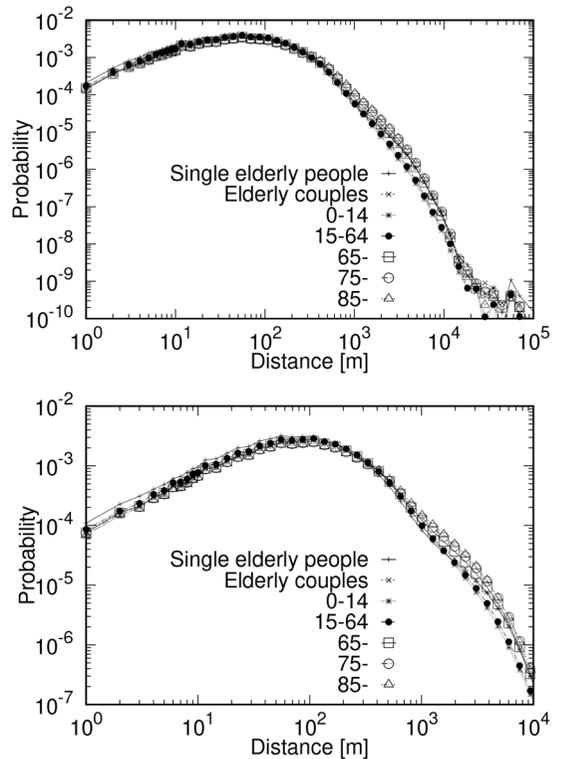


図4 食料品・飲食店 (上図) と病院・医療施設 (下図) について最寄り店舗・施設までの距離の確率密度関数

地が多い一方で、山間部では4km以上、場所によっては8km以上も離れている居住地もある。図3は高知県についての結果である。高知県は山間部が多いため、4km以上離れている居住地が多く存在しており、また、島内に食料品・飲食店がない離島も存在していることが分かる。

次に、全国の国民一人一人について最寄りの食料品・飲食店までの距離を調べた。これは、先程求めた居住地から最寄りの食料品・飲食店までの距離に、その居住地の人口に応じた重みをかけて

⁵⁾ 地球が楕円体であることを考慮して長半径 $a=6378137$, 短半径 $b=6356752.31414$, 第一離心率 $e=\sqrt{\frac{(a^2-b^2)}{a^2}}$ と
 して、経度 x , 緯度 y と経度 x_0 , 緯度 y_0 の二地点間の距離を $a \sqrt{\frac{(y-y_0)^2(1-e^2)^2}{(1-e^2\sin^2(\frac{y+y_0}{2}))^3} + \frac{(x-x_0)^2\cos^2(\frac{y+y_0}{2})}{1-e^2\sin^2(\frac{y+y_0}{2})}}$ に
 より求める。

表1 最寄りの食料品・飲食店までの距離が大きい市区町村

距離 (m)	市区町村名
47791	鹿児島県鹿児島郡十島村
4760	高知県土佐郡大川村
3952	東京都三宅島三宅村
3776	東京都小笠原村
3149	山梨県北都留郡丹波山村
2851	福島県南会津郡檜枝岐村
2837	宮崎県東臼杵郡椎葉村
2714	愛知県北設楽郡豊根村
2700	奈良県吉野郡十津川村
2613	鹿児島県鹿児島郡三島村
2583	福島県双葉郡川内村
2581	東京都御蔵島村
2567	長野県南佐久郡北相木村
2531	宮崎県児湯郡西米良村
2514	宮崎県東臼杵郡諸塚村
2501	福島県双葉郡葛尾村
2490	北海道十勝郡浦幌町
2467	沖縄県国頭郡東村
2386	北海道天塩郡幌延町
2294	熊本県球磨郡球磨村

表2 最寄りの病院・医療施設までの距離が大きい市区町村

距離 (m)	市区町村名
22856	沖縄県島尻郡渡名喜村
7722	岩手県下閉伊郡川井村
6933	沖縄県島尻郡渡嘉敷村
6231	北海道積丹郡積丹町
5928	奈良県吉野郡野迫川村
5362	宮崎県東臼杵郡椎葉村
4846	青森県中津軽郡西目屋村
4472	熊本県球磨郡五木村
4454	北海道阿寒郡鶴居村
4367	北海道宗谷郡猿払村
4341	北海道野付郡別海町
4332	青森県下北郡東通村
4309	青森県下北郡風間浦村
3984	北海道目梨郡羅臼町
3969	大分県東国東郡姫島村
3953	北海道島牧郡島牧村
3952	東京都三宅島三宅村
3942	福島県南会津郡檜枝岐村
3879	福島県双葉郡川内村
3784	宮崎県東臼杵郡諸塚村

集計することで算出される。図4(上図)に得られた距離の確率密度関数を示す。全国民の半数は居住地から約200m以内の距離に食料品・飲食店があり、高齢単身世帯の場合、ほかの世帯に比べて食料品・飲食店の近くに居住している人も多い一方で、数kmも離れている世帯も多い傾向がある。距離を全国民で平均すると279mになるが、65歳以上のみで平均をとると324mと1.16倍大きくなる。したがって、高齢者ほど最寄り食料品・飲食店までの距離が遠い傾向が見られる。同様の傾向は中央値でも確認できる。病院・医療施設についても同様に計算を行った結果を図4(下図)に示す。最寄りの病院・医療施設までの距離の全国民の平均値は432mであるが、65歳以上の平均値は515mと1.19倍大きくなる。したがって、食料品・飲食店と同じく、高齢者の最寄りの病院・医療施設までの距離も遠い傾向が見られる。

国民一人一人について求めた最寄り店舗・施設までの距離を市区町村単位に集計することで、全国の全市区町村の特徴を明らかにすることができる。各市区町村ごとに最寄りの食料品・飲食店、

病院・医療施設までの平均距離を計算した。平均距離の大きい市区町村を表1, 2に示す。距離が大きくなっている鹿児島県十島村、東京都三宅島、東京都小笠原村、東京都御蔵島村、沖縄県島尻郡などの離島では、島内に該当する店舗・施設が電話帳データに存在していなかった可能性が考えられる。また、過疎地域とみなされる高知県土佐郡大川村、山梨県北都留郡丹波山村、岩手県下閉伊郡川井村、北海道積丹郡積丹町などでは、距離が大きくなっている。これらの市区町村ごとに求めた平均距離は、生活環境を定量化する一つの指標として活用できる。

都道府県単位でも同じように調べた。各都道府県ごとに最寄りの食料品・飲食店までの平均距離を計算した結果を図5(上図)に示す。人口の多い都道府県では距離が短く、人口の少ない県では距離が長くなる傾向が見られる。また、65歳以上の人に限定して最寄りの食料品・飲食店までの平均距離を計算した結果を図5(下図)に示す。ほとんどの都道府県で先程の平均距離よりも長くなっており、ここでも高齢者ほど最寄りの店舗までの距

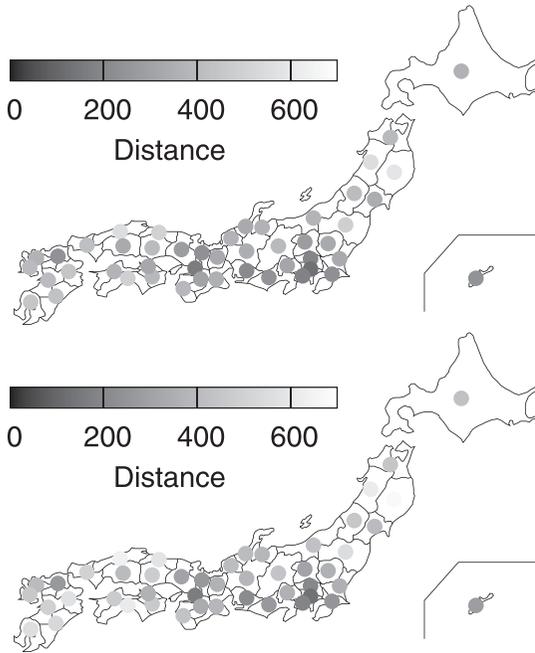


図5 都道府県別に算出した最寄りの食料品・飲食店までの距離 (m) の全年齢平均 (上図) と65歳以上平均 (下図)

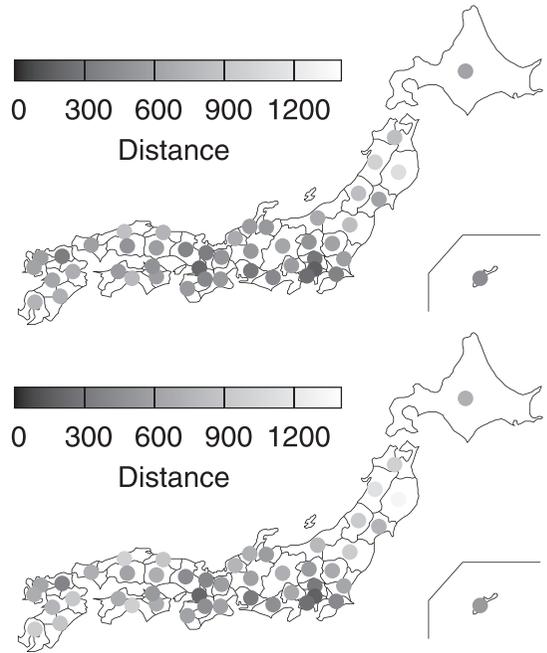


図6 都道府県別に算出した最寄りの病院・医療施設までの距離 (m) の全年齢平均 (上図) と65歳以上平均 (下図)

離が遠い傾向が確認された。ただし、この傾向の強弱は都道府県によって異なり、次節で詳しく調べる。図6は病院・医療施設について計算した結果であるが、食料品・飲食店と同様の傾向が確認できる。

4 最寄り店舗・施設までの距離の比

前節で示したように、高齢者とその他の年齢層では最寄りの店舗・施設までの距離に差異があり、これらの差は市区町村や都道府県によって異なる。この差異を調べるために、各市区町村（都道府県）における65歳以上の最寄り店舗・施設までの平均距離を、全年齢層の平均距離で割った値（距離比）に注目した。この値が大きいほど、高齢者は全住民と比較して最寄り店舗・施設までの距離が遠くなり、高齢者の生活環境が悪いと解釈できる。一般に、移動手段や生活環境は市区町村（都道府県）によって異なるため、距離そのものについては各地域の特性を考慮した上で解釈する必

要があり、地域間の直接的な比較は困難である。しかしながら、距離の比を見ることで、地域の特性を調整した上での評価が可能になる。

食料品・飲食店と病院・医療施設について、市区町村ごとに距離の比を計算した。距離の比の大きい市区町村を表3、4に示す。これらの地域では、高齢者が店舗・施設から相対的に遠くに居住しているため、優先的に施策を検討すべきと考えられる。沖縄県島尻郡南大東村、沖縄県宮古島市、北海道利尻郡利尻富士町では、全住民と比較して高齢者は店舗・施設から約1.4倍も遠くに居住している。距離の比の大きい市区町村は、必ずしも人口の少ない地域や人口減少が顕著な地域とは限らない。例えば、愛媛県松山市や鹿児島県霧島市など、県庁所在地や人口の多い地域も含まれる。買い物や医療へのアクセスに関しては、しばしば人口の少ない地域に焦点が当てられがちだが、実際には高齢者の居住地と店舗・施設の所在地が重要である。両者の距離を考慮して距離の比

表3 最寄りの食料品・飲食店までの距離の比（65歳以上平均/全年齢平均）の大きい市区町村

距離の比	市区町村名
1.40	沖縄県島尻郡南大東村
1.40	沖縄県宮古島市
1.37	北海道苫前郡初山別村
1.36	山口県柳井市
1.36	鹿児島県始良郡始良町
1.35	沖縄県国頭郡今帰仁村
1.35	山口県熊毛郡平生町
1.35	長野県北安曇郡白馬村
1.33	鹿児島県大島郡瀬戸内町
1.30	宮崎県日向市
1.30	和歌山県有田郡有田川町
1.28	愛媛県松山市
1.28	福岡県筑紫郡那珂川町
1.27	大分県中津市
1.27	広島県三次市
1.27	和歌山県田辺市
1.27	京都府福知山市
1.27	鹿児島県霧島市
1.26	沖縄県国頭郡国頭村
1.25	和歌山県伊都郡高野町

表4 最寄りの病院・医療施設までの距離の比（65歳以上平均/全年齢平均）の大きい市区町村

距離の比	市区町村名
1.38	北海道利尻郡利尻富士町
1.37	山口県熊毛郡平生町
1.37	鹿児島県霧島市
1.37	和歌山県有田郡有田川町
1.35	沖縄県宮古島市
1.34	北海道北斗市
1.34	岐阜県本巣市
1.33	青森県むつ市
1.33	沖縄県島尻郡南大東村
1.33	北海道中川郡中川町
1.32	北海道虻田郡喜茂別町
1.32	沖縄県国頭郡国頭村
1.29	和歌山県日高郡日高町
1.29	京都府福知山市
1.29	福岡県筑紫郡那珂川町
1.29	鹿児島県始良郡始良町
1.28	沖縄県うるま市
1.28	鹿児島県大島郡瀬戸内町
1.28	宮崎県日向市
1.27	北海道標津郡中標津町

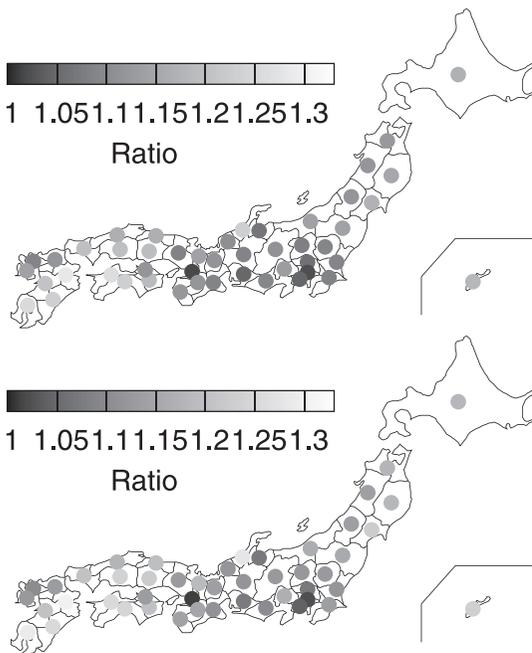


図7 食料品・飲食店（上図）と病院・医療施設（下図）について都道府県別に算出した最寄り店舗・施設までの距離の比（65歳以上平均/全年齢平均）

を算出したことで、より実態に即した状況把握が可能になった。

都道府県単位で距離の比を求めた結果を図7に示す。食料品・飲食店と病院・医療施設のいずれにおいても、東京都、大阪府、神奈川県では高齢者とほかの年齢層で大きな違いは見られなかった。しかし、大分県、鹿児島県、愛媛県、石川県、宮崎県、高知県では、高齢者が店舗・施設から遠くに住んでいる傾向が強く現れている。地域包括ケアなどの行政政策を検討する際には、地域特性を十分に考慮することが重要であり、これらの知見の活用が期待できる。

5 最寄り店舗・施設までの距離の変化率

以上は一時点のデータについての分析であったが、本節では時間的な変化に着目した分析を行う。具体的には、2010年と2015年の2つの時点と比較し、最寄りの店舗や施設までの距離がどのように変化したかを調べた。その際、2015年の距離の中央値 Y が、2010年時点の距離の中央値 X に対して何倍になったかを表わす変化率 $R = Y/X$ に注

表5 食料品・飲食店についての距離の変化率の比の大きい市区町村

市区町村名	距離の変化率の比
鹿児島県鹿児島郡三島村	2.34
鹿児島県大島郡大和村	1.51
北海道阿寒郡鶴居村	1.49
福島県双葉郡広野町	1.34
長野県下伊那郡根羽村	1.30
北海道古宇郡神恵内村	1.28
山形県最上郡大蔵村	1.26
奈良県吉野郡上北山村	1.24
青森県三戸郡新郷村	1.22
兵庫県赤穂郡上郡町	1.20
沖縄県国頭郡大宜味村	1.20
福島県大沼郡金山町	1.19
沖縄県中頭郡中城村	1.19
愛知県知多郡美浜町	1.19
愛媛県上浮穴郡久万高原町	1.18
富山県中新川郡舟橋村	1.17
福島県河沼郡会津坂下町	1.17
千葉県香取郡神崎町	1.17
北海道河東郡士幌町	1.15
北海道上川郡清水町	1.14

表6 病院・医療施設についての距離の変化率の比の大きい市区町村

市区町村名	距離の変化率の比
岡山県勝田郡勝央町	1.60
北海道虻田郡ニセコ町	1.47
島根県隠岐郡西ノ島町	1.44
福島県双葉郡広野町	1.41
奈良県吉野郡野迫川村	1.37
北海道様似郡様似町	1.30
新潟県岩船郡粟島浦村	1.29
北海道新冠郡新冠町	1.25
北海道河東郡士幌町	1.25
鹿児島県大島郡喜界町	1.24
山口県阿武郡阿武町	1.240
奈良県吉野郡上北山村	1.24
沖縄県島尻郡伊是名村	1.20
長野県小県郡長和町	1.20
北海道上川郡東川町	1.20
新潟県南魚沼郡湯沢町	1.19
岐阜県大野郡白川村	1.19
長野県下伊那郡壳木村	1.17
福島県耶麻郡北塩原村	1.16
長野県上水内郡飯綱町	1.15

目した。全国の住民を対象にして調べた結果、変化率 $R_{住民}$ は食料品・飲食店では0.945、病院・医療施設では0.972であり、5年間で距離が短縮されたことが明らかになった。一方、65歳以上のみで調べた場合、変化率 $R_{高齢者}$ は食料品・飲食店で0.956、病院・医療施設で0.984であった。高齢者の変化率はほかの年代と比較して1に近く、相対的に最寄りの店舗・施設までの距離が遠くなっており、高齢者の生活環境の悪化が懸念される。

なお、変化率が1より小さいからと言って、実際に5年間で距離が短くなったと断言することはできないことに注意する必要がある。異なる二つの時点のデータを用いて分析しているが、各時点のデータでサンプルバイアスの入り方が同じである保証はない。実際、2010年と2015年で国勢調査の方法はまったく同じではなく、固定電話の利用が減少していることから電話帳に記載されている店舗・施設が減少している可能性も考えられる。一方、全国民で求めた変化率 $R_{住民}$ と高齢者で求めた変化率 $R_{高齢者}$ を比較することで、両者は同じ条件で算出したものであるためバイアスの影響が軽

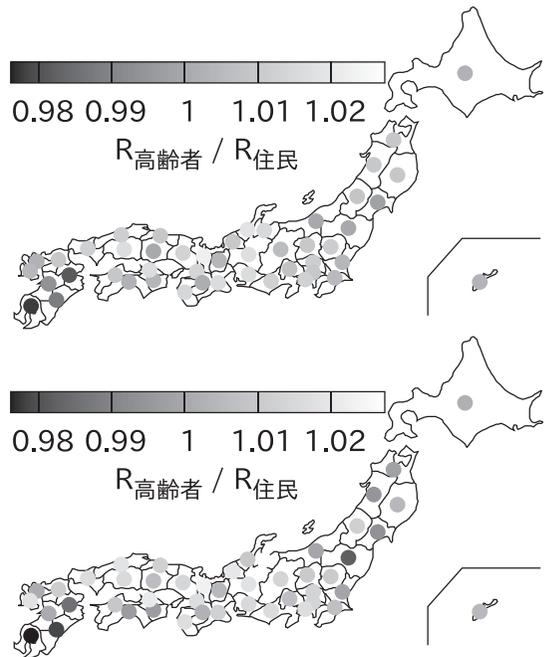


図8 食料品・飲食店（上図）と病院・医療施設（下図）について都道府県別に算出した距離の変化率の比

減され、信頼性の高い評価が得られると言える。

そこで、高齢者と全住民の変化率の比 $R_{\text{高齢者}}/R_{\text{住民}}$ に注目した。この比が1より大きい(小さい)ほど、5年間で高齢者の最寄り店舗・施設までの距離が相対的に遠く(近く)なったことを示す。市区町村単位で計算した結果、変化率の比が大きかった市区町村を表5、6に示す。これらの地域では高齢者の生活環境が相対的に悪化しており、特に高齢化や人口減少の課題を抱えている山村地域や離島地域でその傾向が顕著であることが分かる。同様に都道府県単位で計算した結果を図8に示す。鹿児島県、大分県、宮崎県では変化率の比は1より小さく、相対的にみて高齢者の最寄り店舗・施設までの距離が短縮されたことが分かる。一方、京都府や香川県では変化率の比は1より大きく、距離が相対的に長くなったことが分かる。このように高齢者と全国民の変化率の比を観測することで、最寄り店舗・施設までの距離の側面から高齢者の生活環境の悪化、改善の状況を定量化することができた。

6 まとめと今後の課題

本研究では、全国の住民一人一人に対し、100mメッシュの精度で食料品・飲食店や病院・医療施設までの距離を計算した。食料品や医療施設へのアクセスに必要な距離を100mの空間解像度で可視化することで、どの地域にどの程度の問題が生じているかを全国規模で詳細に把握することができるようになった。さらに、市区町村や都道府県の単位で集計を行い、各地域の特徴を明らかにした。各地域ごとに高齢者と全住民を比較することで、高齢者の生活環境を評価し、悪化や改善の傾向を把握することができた。これらの知見を用いて、買い物や医療へのアクセスが困難であると疑われる地域を早期に特定し、実態を把握するためにその地域に限定してアンケート調査を行うなど、ビッグデータとアンケート調査を組み合わせた効果的な活用が考えられる。

本研究では、各地域を平均距離で特徴づけて議論を行った。同じ地域に住む住民でも、距離の遠い人もいれば、近い人もおり、図4で見たように距

離は広く分布している。今後、各地域ごとに距離の分布をさらに調べていくことで、人によって距離の差が大きい地域を特定したり、全国で距離の分布が類似した地域をクラスタリングすることが可能である。本研究では、2010年と2015年の二時点について分析を行ったが、国勢調査は5年に一度、電話帳データは数ヶ月ごとに更新されるため、過去のデータを活用して例えば1年ごとに分析することも可能である。これにより、時系列として各地域の変化の推移を追跡し、現在の状況に至った経緯を調べることもできる。このような分析によって得られる結果は、各地域の生活環境の評価や政策立案に有益なものになると期待される。

今後の課題として、以下のことが考えられる。まず、国勢調査に回答していない住民、電話帳に記載されていない店舗や施設の存在が挙げられる。これらの情報の欠落によるサンプルバイアスの影響を確認し、実際の状況を反映しているかどうかを精査する必要がある。特に、農村部では宅配や移動販売、自家生産、直売所などの特殊なサービスや施設が存在することを考慮しなくてはならない。算出された指標が実際の状況を正確に反映しているかどうかの調査も必要である。例えば、本研究では直線距離として距離を推定したが、精度を向上させるために道路に沿った距離を用いることもできる。特に、川や海、山などの地形がある場合は、この誤差の影響は大きいと考えられる。経路、地形の起伏、移動手段、年齢なども考慮して、より現実に応じた指標を検討することも考えられる。また、今回の分析で得られた変化の原因が店舗や施設の撤退や進出によるものなのか、それとも住民の転居、死亡、出生によるものなのかを検討する必要もある。実際に施策や政策を策定する際には、データのみからでは把握できていないこれらの要因も慎重に考慮した上で行う必要がある。

III 保険医療ビッグデータ活用の可能性

医療保険のあらゆる領域において、ビッグデー

タの活用が期待される。医療画像、医療記録、ゲノムデータ、創薬などでは、より効率的な診断、治療、予防に向けた取り組みが行われている。今後も情報通信技術が急速に進展することが想定されるため、ウェアラブルデバイスやセンサー技術などの活用が進み、心拍数や血圧などのさまざまな生体データの取得がより容易になり、日常生活でのあらゆるデータが収集可能になると予想される。世界的に見れば、日本は超高齢化社会であるため、特に高齢者に関するデータは今後益々取得できるようになると考えられ、これを活用した研究の進展が望まれる。一方で、個人情報保護と倫理的観点から、データの適切な管理と利用がより求められるようになる。個人情報の保護が過度に強調されると、データの有効な活用が制限されてしまう可能性があるため、このバランスを取ることが重要になってくる。データを適切に分析するためには、その領域の専門知識だけでなく、ビッグデータの取り扱いに関する情報と数理の知識やスキルも必要になる。1人の研究者が両方の知識やスキルを修得するのは難しいため、専門知識を持った複数の人材が密に連携して分析を進めることが今後より重要になると考えられる。

謝辞

本研究は東大CSIS共同研究No.1306の成果の一部である（座標付き電話帳DBテレポイント法人版、ゼンリン提供）。

参考文献

- 大西立顕 (2021) 「離散幾何学とその周辺」, 『数理科学 2021年12月号』, Vol.59, no.12, pp.66-72。
 藤原義久 (2024) 「経済現象とネットワークの科学」, 『日本物理学会誌』, Vol.79, no.1, pp.4-11。
 大西立顕・渡辺努 (2019) 「経済バブルの数理モデリング」, 『数理科学 2019年6月号』, Vol.57, no.6, pp.15-21。
 Ito, Mariko I. and Takaaki Ohnishi (2021) “Co-occurrence Network of TV Advertisements Revealing Japanese Lifestyle,” *Applied Network Science*, Vol.6, No.1, pp.1-19.

- Mizuno, Takayuki, Takaaki Ohnishi and Tsutomu Watanabe (2021) “Visualizing Social and Behavior Change due to the Outbreak of COVID-19 Using Mobile Phone Location Data,” *New Generation Computing*, Vol.39, pp.453-468。
 Saito, Saki, Mariko I. Ito and Takaaki Ohnishi (2022) “Fluctuations in the Number of Stores by Industry During the COVID-19 Pandemic Based on Japanese Phone Book Entries,” *The Review of Socionetwork Strategies*, Vol.16, pp.545-557。
 経済産業省 (2015) 『買物弱者・フードデザート問題等の現状及び今後の対策のあり方に関する調査報告書』, <http://www.meti.go.jp/policy/economy/distribution/kaimonojakusyashien.html> (2024年2月14日アクセス)。
 農林水産政策研究所 (2018) 『食料品アクセスマップ』, http://www.maff.go.jp/primaff/seika/fsc/faccess/a_map.html (2024年2月14日アクセス)。
 岩間信之 (2013) 『フードデザート問題無縁社会が生む「食の砂漠」』, 農林統計協会。
 ——— (2017) 『都市のフードデザート問題：ソーシャル・キャピタルの低下が招く街なかの「食の砂漠」』, 農林統計協会。
 薬師寺哲郎 (2015) 『超高齢社会における食料品アクセス問題—買い物難民, 買い物弱者, フードデザート問題の解決に向けて—』, ハーベスト社。
 三浦英俊・古藤浩 (2010) 「メッシュデータを用いた人口減少地域における買い物距離の分析」, 『都市計画論文集』, Vol.45, no.3, pp.643-648。
 鳥海重喜 (2014) 「福岡市におけるフードデザート問題の分析」, 『都市計画論文集』, Vol.49, no.3, pp.993-998。
 中村みず季・渡辺理絵 (2014) 「フードデザートマップを用いた後期高齢者の買い物環境」, 『地理空間』, Vol.7, no.1, pp.33-50。
 中村恵美・浅見泰司 (2019) 「経済的アクセス困難性からみた大都市中心部におけるフードデザート問題の実態把握と規定要因」, 『日本建築学会計画系論文集』, Vol.84, no.756, pp.437-445。
 衣笠匠斗・樋野公宏・別所あかね・貞広幸雄 (2023) 「集約型都市政策下での食料品店舗への近接性の時系列変化」, 『都市計画論文集』, Vol.58, no.3, pp.735-742。
 岩間信之・今井具子・田中耕市・浅川達人・佐々木緑・駒木伸比古・池田真志 (2018) 「食料品充足率を加味した食料品アクセスマップの開発」, 『フードシステム研究』, Vol.25, no.3, pp.81-96。

(おおにし・たかあき)

The application of data science to healthcare

OHNISHI Takaaki*

Abstract

The depopulation and labor shortage in Japan have made it increasingly difficult for residents to access shopping and medical services throughout the country. As an application of data science to healthcare, this study estimated the distance from each individual's location to the nearest stores and facilities using national census data and geocoded corporate phone directory data at a 100-meter mesh scale. When comparing data from 2010 to 2015, it was found that the distance decreased for all age groups over the five-year period, except for the elderly who experienced a relative increase in distance. The estimated distances were aggregated by prefecture and municipality, stratified by age groups, and used to identify areas where access to shopping and medical services is challenging. Additionally, the study evaluated the living environment of the elderly individuals and identified trends in its deterioration or improvement.

Keywords : Big geospatial data, Phone directory data, Shopping difficulties, Medical difficulties, Nearest stores/ facilities

* Professor, Graduate School of Artificial Intelligence and Science, Rikkyo University