公的統計に対する差分プライバシーの適用と有効性の評価に関する検討―国勢調査を例に―

伊藤 伸介†
寺田 雅之††
加藤 駿典†

Application of Differential Privacy and Assessment of Its Effectiveness for Official Statistics: Examples from the Japanese Population Census

ITO Shinsuke
TERADA Masayuki
KATO Shunsuke

海外の統計作成部局では、近年公的統計におけるプライバシー保護のために攪乱的手法が採用されてきた。アメリカセンサス局は、複数の公表された人口センサスの統計表の組み合わせによって個体情報を特定しようとする再構築攻撃への対応策として、2020年人口センサスを対象に、差分プライバシーの実現方式が適用された統計表の作成・公表を行った。それに関しては、データの利用者や利害関係者が要求する統計表の結果数値の精度も考慮した上で、差分プライバシーの実用性が追究された。本稿では、このようなアメリカにおけるプライバシー保護における攪乱的手法の適用に関する動向を踏まえつつ、国勢調査の個票データから作成された統計表に適用した、各種の差分プライバシーの実現方式について比較・検証を行う。また、差分プライバシーの有効性に関するこれまでの実証研究をもとに、公的統計における差分プライバシーの適用に関する方法的な可能性についても論じる。

キーワード: 差分プライバシー、2020年人口センサス、一般公開型ミクロデータ、国勢調査、トップダウン法、生態学的誤謬

A growing number of National Statistical Institutes in various countries have adopted perturbative methods as a mechanism for privacy protection for official statistics. The U.S. Census Bureau has released statistical tables for the 2020 Population Census that use differential privacy was to minimize identification risks from reconstruction attacks via the combination of published statistical tables. The U.S. Census Bureau has also investigated the practicality of differential privacy, ensuring the accuracy for cells demanded by data users and stakeholders. This paper conducts comparative research into the effectiveness of differential privacy for statistical tables created using individual data from the Japanese Population Census data to which we apply perturbative methods similar to the methods for official statistics in the United States. This paper also discusses the potential of differential privacy as a methodology for Japanese official statistics based on previous empirical research on the effectiveness of differential privacy.

Keywords: differential privacy, 2020 U.S. Census, public use microdata, Japanese Population Census, top-down data construction method, ecological fallacy

† 中央大学経済学部 Email:ssitoh@tamacc.chuo-u.ac.jp
†† (株)NTT ドコモ Email: teradam@nttdocomo.com
† 総務省統計局 Email: shun911k@gmail.com

1. はじめに

公的統計におけるプライバシー保護のための秘匿技法として、海外の統計作成部局は、攪乱的手法(perturbative methods)を積極的に採用してきた(伊藤・寺田(2023))。特に、アメリカセンサス局(以下「センサス局」と呼称)は、コンピュータサイエンスの分野で展開されてきた、差分プライバシー(differential privacy)の方法論に基づく攪乱的手法の公的統計への適用可能性を模索してきた。センサス局は、アメリカにおいて 2020 年に実施された人口センサス(以下、「センサス」の略称)において、差分プライバシーを用いた統計表の作成の検討を進めてきた。特に小地域レベルのセンサスの統計表を作成する場合に、複数の公表されたセンサスの統計表の組み合わせによって個体情報を特定しようとする再構築攻撃(reconstruction attack)への対応として、センサスを対象にした差分プライバシーの方法論の実用性が模索されてきた(伊藤・寺田(2020), 伊藤他(2022))。その結果、センサス局は、2020 年センサスを対象に、差分プライバシーの実現方式が適用された統計表を作成・公表した。

統計実務における差分プライバシーの適用にあたっては、主として人口センサスに関する統計表の作成・公表に焦点が当てられている。しかしながら、公的統計を対象にした差分プライバシーについての最近の研究動向をみると、合成データ(synthetic data)の作成に対する潜在的な可能性が指摘できる[1]。こうしたことから、わが国でも差分プライバシーの公的統計への適用可能性を追究することは、わが国において将来的な公的統計の統計表の作成・提供や公的統計の二次利用の将来的な方向性を議論する上でも、検討対象に値すると思われる(伊藤(2022))。

本稿では、最初にアメリカを対象に、プライバシー保護における攪乱的手法の適用に関する動向を概括した上で、差分プライバシーの方法論が適用された統計数値における秘匿性と有用性に関する定量的な評価について論じる。つぎに、わが国の国勢調査の個票データから作成された統計表に各種の差分プライバシーの実現方式を適用した上で、評価指標をもとに有用性についての比較・検証を行う。さらに、差分プライバシーの有効性に関する先行研究をもとに、公的統計における差分プライバシーの適用に関する方法的な可能性についての試論的な考察を行う。これらの論考を踏まえ、最後に、国勢調査に対する差分プライバシーの方法論の適用可能性について議論する。

2. アメリカにおけるセンサスデータの秘匿措置と定量的な評価をめぐって[2]

センサス局は、「フォーマルな(formal)」プライバシーの１つである差分プライバシーの方法論のセンサスへの適用可能性を追究してきた。差分プライバシーは、未知の攻撃を含めた任意の攻撃に対する「包括的 (ad omnia)」な安全性 (Dwork (2007)) を実現することを目的としたプライバシー保護の枠組みであり、様々なプライバシー保護手法に対して統一的な安全

[1] 例えば、2023 年 9 月にドイツのヴィースバーデンで開催された、UNECE(United Nations Economic Commission for Europe、国連欧州経済委員会)主催の「統計データの機密保護に関する専門家会議(UNECE Expert Meeting on Statistical Data Confidentiality)」においては、差分プライベートなモデル等から生成された合成データについて比較・検証を行った事例がある。

[2] 本節の執筆においては、下記のセンサス局の 2020 年センサスに関するウェブサイトに掲載されている人口センサスに関するデータの公表や秘匿措置に関する資料を参照した。
https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-development.html

性指標を定量的に与える[3]。この指標は $\varepsilon\ (\geq 0)$ で表され、その値が小さいほど安全性が高いことを示す。なお、このεは、プライバシー損失予算(privacy loss budget)とも呼ばれる。

差分プライバシーの方法論の 2020 年センサスへの適用にあたっては、統計データの作成者側であるセンサス局と利用者側である研究者の両方から、差分プライバシーの是非について様々な意見が出された。また、差分プライバシーの安全性指標であるパラメータεにおいては、統計データの作成・公表をめぐる「ステークホルダープロセス」の中で、εの値が決定された(伊藤・寺田(2023))。このεの設定の過程で、センサス局は、差分プライバシーの方法論が適用された攪乱済のセンサスデータにおける結果数値と元データでの数値から誤差を攪乱済センサスデータの有用性の指標として公表してきた(伊藤他(2022))。このように差分プライバシー適用については、その実現方式に関する定量的な評価も求められる。

本節では、最初にアメリカにおける差分プライバシーの適用の経緯やそれに関するセンサス局の議論について概括する。つぎに、差分プライバシーの実現方式の有効性を比較・検証した Christ らの実証研究(Christ et al. (2022))を考察の素材として、攪乱済のセンサスデータの秘匿性と有用性に関する評価方法について述べる。

2.1 センサスミクロデータの作成・提供状況

センサスにおける差分プライバシーの導入の経緯を確認するために、本節では、センサスデータ、特に統計表の元データとなるミクロデータの作成・提供状況を概観する。

センサス局が、センサスを対象にミクロデータの公開を開始したのは、1960 年センサスからであった。1962 年に、センサスの詳細(long form)調査票の対象世帯(全世帯の約 1/4)に対してサンプリングが施されるだけでなく、直接的な識別子の削除や人口規模が 25 万人以上の地域区分の設定によって、全人口の 0.1%のレコードが含まれるセンサス最初の一般公開型ミクロデータサンプル(Public Use Microdata Sample=PUMS)が作成・提供された(McKenna(2019))。その後、1970 年から 2010 年までのセンサスにおいては、PUMS の作成のために、サンプリング、トップコーディング、リコーディングといった各種の非攪乱的手法が適用された。その時期のセンサスにおける具体的な匿名化技法として、1980 年センサスから、複数ファイルを作成するために、世帯と個人の抽出単位に合わせて、1%と 5%という異なるサンプリング率が設定された。また、地域区分の閾値が 10 万人以上に縮小された。1990 年センサス以降においては、トップコーディング(全体の分布の上位 0.5%かあるいは非負である分布の上位 3%)およびリコーディングの適用(居住地、勤務地、集合住宅 (group quarters)、所得、年齢等)、欠測値の補完が行われた。さらに、2000 年センサスでは、カテゴリカルな変数における分類区分のしきい値が 1 万人以上に設定された。

1990 年センサスまでは、PUMS の作成のために、非攪乱的手法のみが適用されてきた。2000 年以降のセンサスでは、ミクロデータに対して攪乱的手法が新たに採用されたのが、重要な特徴だと言える。2000 年と 2010 年のセンサスでは、公表統計表の元になる個票データにスワッピングが適用された。同様に、2000 年センサスから、世帯人員の多い世帯にノイズ付与(noise addition)が行われただけでなく、2010 年センサスでは、集合住宅について一部合成データの方法論の適用も追加的に適用された。このようにセンサス局は、2010 年センサスまでは、「その場限りの (ad hoc)」秘匿措置によって、PUMS の作成・公開を行ってきた(McKenna(2019b), 伊藤他(2022))[4]。

---

[3] 決定論的手法に対しては $\varepsilon \to \infty$ となる (安全性が与えられない)。

[4] 2000 年と 2010 年のセンサスで用いられてきたスワッピング技法については、ターゲット・スワッピングが、

それに対して、センサス局は、スワッピングが適用されたセンサスデータでは、再構築攻撃への対応が困難であるという理由から、「フォーマルな」プライバシーである差分プライバシーの方法論の実現可能性を模索し、2020 年センサスの公表統計表における露見制御のための新たな方法論として、差分プライバシーの実現方式を採用した。なおセンサス局の担当者によれば、2020 年センサスについては、サンプルデータではなく、全レコードを対象にしたPUMF(=Public Use Microdata File)として作成し、2024 年に PUMF を提供することを計画しているとのことである。さらに、センサス局が、差分プライベートな(differential private)ミクロデータとして、PUMF を作成・提供する予定であることは、特筆すべき点であると言えよう。

2.2 2020 年センサスにおける差分プライバシーに向けた議論

センサス局は、2010 年センサス以前からプライバシーに関する懸念を持っていた。それが、2001 年にセンサス局内部に DSEP(=Data Stewardship Executive Policy Committee)が設置される契機となった(McKenna(2019a))。なお、DSEP は、センサス局がアメリカ国民及びアメリカ経済に関するデータを効率的かつ合理的に収集・利用することを保証することをその任務としていることから、データの収集過程全体にわたって回答者のプライバシーとデータの秘匿情報を保護することが求められる。

センサス局は、2000 年代～2010 年代にかけて、センサス局が一般に公開した「所得・プログラム参加調査(Survey of Income and Program Participation)」の Public Use File(PUF)や American Community Survey の Public Use Microdata Sample(PUMS)と外部の公開データとのマッチングによる再識別に関する実験を行ってきた。その結果、PUF や PUMS について露見リスクがあることが明らかにされた(McKenna(2019a))。そこで、2020 年センサスについては、DSEP は、2010 年センサスまで適用されたスワッピングとは異なる新たな方法論の適用の方向性を模索した結果、差分プライバシーの方法論の可能性を検討するに至った(Jamin(2021))。

センサス局は、2020 年センサスの公表統計表において差分プライバシーの方法論を適用するために、2010 年センサスを用いた差分プライバシーの実用性に関する検証を行った。具体的には、統計表の公表によって消費されるプライバシー損失予算 ε を設定し、TopDown アルゴリズムと呼ばれる地域のレベルにおけるプライバシー損失予算の割り当てについての検証を進めてきた(Garfinkel et al.(2019), 伊藤・寺田(2020))。これは、再構築攻撃へのセンサス局の対応と見ることができる。

差分プライバシーの適用におけるパラメータ ε の設定に関しては、統計数値の秘匿性の観点だけでなく、データの利用者や利害関係者が要求する統計表の結果数値の精度も考慮するステークホルダープロセスのもとで、DSEP での議論等に基づいて、数度の修正がなされた。ステークホルダープロセスは、2010 年センサスの個票データに対して「露見回避システム(Disclosure Avoidance System=DAS)を適用したテストデータ (demonstration data) を作成・公開し、テストデータの利用者からのフィードバックを 2020 年の DAS に反映するというプロセスを繰り返すことによって実施された。

差分プライバシーが適用された統計表は、区画改定データ(redistricting data)として公表されることから、調査時点で実査によって捕捉された数値と誤差が、5%以内に収まっていることが目標精度となった[5]。したがって、州や郡レベルの人口、選挙年齢人口、アメリカ行政管理

欧州委員会の勧告を受けて、イギリスやドイツのようなヨーロッパ諸国でも、センサスを対象とした秘匿技法として用いられた実績があることが知られている。

[5] センサス局は、選挙区改正用のデータにおいて精度評価に関する目標値の計測のために、司法省から選挙区の改正に関する利用のサンプルを提供されている。

予算局(Office of Management and Budget=OMB)によって選定された人種・民族グループが占める比率といった数値は、選挙区の改正と投票権法の執行のために信頼できる数値であることが追究された[6]。さらに、DSEP は、差分プライバシーの実現方式の適用において、アメリカインディアン/アラスカ先住民/ハワイ先住民(American Indian/Alaska Native/Native Hawaiian=AIANNH)の地域および準州地域の人口数とその人口社会的特性を直接算出可能なように、TopDown アルゴリズムの修正を図る必要があった。

これらのプロセスに伴う数度の ε の修正の後、2021 年 6 月に公開された最終版のプライバシー保護済ミクロデータファイル(Privacy-Protected Microdata Files=PPMFs)の作成においては、DSEP において、DAS における全体のプライバシー損失予算のパラメータが ε=19.61 と設定された(伊藤他(2022))。このパラメータの数値をもとに、2021 年 8 月 16 日に、2020 年センサスの区画改定データ(PL94-171) がセンサス局で公表された。なお、センサス局の担当者によれば、次回の 2030 年センサスにおいても、差分プライバシーの実現手法を適用することを計画しているとのことである。

2.3 スワッピング技法と差分プライバシーの実現方式の有効性に関する比較・検証について

本節では、スワッピング技法と差分プライバシーの方法論の有効性についての定量的な比較・検証を指向した、センサスを用いた実証研究の事例として、Christ らによる 2010 年のセンサスデータを用いた研究成果を紹介したい。Christ et al. (2022)は、最初に 2010 年センサスの公表されたセンサスブロックや群レベルのデータを用いて、より大規模な 600 人～1500 人の住民を対象としたブロックグループ(block group)から構成される実験用の合成データを作成した。また、公表されているセンサスブロックレベルで統合することによって、合成データには、ブロックグループごとに性別、年齢だけでなく人種、ヒスパニックの有無、世帯の規模、住宅の所有形態(household tenure)に関する変数も組み込まれている。

このように作成された検証用のミクロレベルのセンサスデータに対して、本研究では、各種のスワッピング技法と差分プライバシーの実現方式が適用された。前者については、特定のスワッピング率を設定した上で、特定の州あるいは郡に関するレコード群を対象に、州・郡における人口分布に基づいて、異なるセンサスブロック同士でスワッピングが実行される。具体的には、本実験では、①真の(true)ランダム・スワッピング、②擬似的な(pseudo)ランダム・スワッピング、③類似レコードスワッピングが用いられている。真の(true)ランダム・スワッピングでは、州に関する真の分布からスワッピングの対象レコードがランダムに選ばれて、レコードの入れ替えが行われる。②擬似的な(pseudo)ランダム・スワッピングの場合には、州のセンサスデータの中で年齢と性別について一致するサブデータセット同士で入れ替えが行われる。したがって、年齢と性別については真の値が保持される。その上で、人種については、州の分布からランダムな抽出がなされる。③類似レコードスワッピング(similar swapping)においては、類似性の閾値を設定した上で、人種、年齢と性別といった変数において閾値の基準を満たすように、入れ替えが行われる。具体的には、年齢に関しては年齢階級区分、人種については分類区分における近接するカテゴリーが類似性に関する閾値の設定対象となっている。

---

[6] 選挙区の改正と投票権法の執行にあたっては、議会地区(congressional districts)のような行政区域の人口が必要であるが、議会地区は人口数が公表されなければ区域を定められないため、事前に設定することはできなかった。そこでセンサス局は、精度評価を行うために、選挙区の構成と規模が類似した、すでに設定されている地域から構成された３つの区域(specified geographic constructs)を用いた。具体的には、①ブロック群(TopDown アルゴリズムの対象となった地理的階層単位"spine")、②センサスの実査のために設定された区域("off-spine")、③小行政区画(minor civil division)のある州とそれ以外の州を区別する目的で編成された"off-spine"が用いられている。

差分プライバシーの実現方式に関しては、年齢階級区分、ヒスパニックの有無(2 区分)、人種(63 区分)、世帯規模(4 区分)でクロスされるセルを対象に、センサス局が実施した Top Down アルゴリズムが本実験でも適用されている。それによって、クロスされた各セルの度数にノイズを付加した出力結果が得られる。また、州レベルの人口数が公表数値と正確に一致するだけでなく、負の値については 0 に変換するといった処理も考慮した上で、ノイズが付与されている。

つぎに、有用性と秘匿性の評価方法について述べる。Christ et al. (2022)は、スワッピングや差分プライバシーの実現手法によって攪乱が施されたデータの元データに対する有用性に関する評価指標として、対象となる変数のヒストグラムについて平均平方誤差(MSE)および平滑化されたカルバック・ライブラー情報量(μ-smoothed Kullback-Leibler divergence、以下「KL 情報量」と呼称)を用いている。スワッピング済データについてはスワッピング率を変更した場合、差分プライバシー適用済みデータに関してはパラメータ ε を変えた場合の KL 情報量や MSE を見ていくと、スワッピング率が減少およびパラメータ ε が上昇するにしたがって、MSE および KL 情報量が小さくなっており、攪乱が施されたデータが元データに対して相対的に近似する傾向にあることが確認できる。したがって、スワッピングと差分プライバシーのいずれも、統計数値の精度については比較可能であることがわかる。Christ et al. (2022)によれば、興味深い結果として、スワッピングを適用した場合、カリフォルニア州アラメダ(Alameda) 郡のような多人種の市民が居住している地域では、オレゴン州ワシントン(Washington)郡のような少数の人種(この場合白人)が住民の大半である郡と比較して、KL 情報量が大きくなる傾向にある。それに対して、差分プライバシーを適用しても、人種の分布が異なる２つの郡における KL 情報量の推移に大きな違いが見られないことが確認される。

本実験では、秘匿性の評価指標に関しては、スワッピング済みのデータを対象に一般に公表されている外部情報としてのデータセットとのリンケージを行い、スワッピングされたレコードが特定化されるリスクの定量的な評価を行っている。具体的には、年齢階級区分、センサスにおける人種の区分と世帯規模を用いたマッチングを行い、照合されたレコードの組においてヒスパニックの有無に関する属性値も一致している場合に、スワッピング済みのデータに識別可能あるいは識別されたと判断することができる。本実験においては、全般的にスワッピング率が上昇するにしたがって、スワッピング済データが外部情報とのリンケージによって、個体情報が特定されるリスクが傾向的に下がっていることが確認できる[7]。なお、多人種の郡において、マイノリティの人種が居住している地域に限定すると、スワッピング率を増大した場合でも(スワッピング率を 100% にした場合においても)個体の識別リスクが相対的に高くなる傾向にあることがわかる。

なお、Christ et al. (2022)では、差分プライバシー適用済みデータについては、クロス集計された各セルに対してノイズ付加を行うことから、スワッピングと同様にレコードリンケージによる攻撃で秘匿性の検証を試みようとしても、レコードに個体識別子が付与されていないため、リスク評価を行うことは困難であると判断されている。

[7] センサスデータを含む人口社会データにおいて、再識別の確率に関する一般的な基準は存在しないと考えられる。それに対して、Christ et al.(2022)によれば、医療データに関しては、医療情報のプライバシー保護に関する法律である HIPPA(=Health Insurance Portability and Accountability Act)の Safe Harbor Law に基づき、医療データの安全性を判断するための「国民が許容可能な再識別リスクに関する基準(nationally accepted standard of re-identification risk)」が考えられており、その基準値は、0.04%であることが指摘されている。本実験における結果に基準値である 0.04%を適用した場合、多人種の市民が居住しているアラメダ郡においては、スワッピング率を 100%(1.0)にしても、個体が識別される確率は少ないケースでも 4%となっていることから、許容可能な再識別リスクに関する基準値を大きく上回っていることが確認できる。

こうしたスワッピングと差分プライバシーの比較研究の成果を要約すると、以下のとおりである。第1は、パラメータεの値にもよるが、差分プライバシーの実現手法が適用されたデータにおいても、スワッピングを適用した場合と同等のデータの精度を担保することが可能なことである。第2は、スワッピングにおいては、人種に着目した場合、多人種の市民が居住する地域、とくにマイノリティの居住地域においては、スワッピング率を上昇させても、個体識別リスクが増大傾向にあることである。

3. 国勢調査への差分プライバシーの適用に関する諸課題

差分プライバシーを実現するメカニズムは様々な種類のものが存在し、対象となるデータの性質や、抽出したい出力データの種類 (問合せの種類) などによって、それぞれ適・不適がある。つまり、わが国の国勢調査の統計表に対してすべてのメカニズムが適用可能であるとは限らない。また、たとえ適用すること自体は可能だったとしても、不適切なメカニズムを用いると出力される統計の有用性は大きく損なわれる。これらを踏まえ、本節では、わが国の国勢調査に対して差分プライバシーの方法論を適用するにあたって考慮すべき課題を論じる。

3.1 国勢調査への差分プライバシーの適用にあたっての課題

差分プライバシーを満たす代表的なメカニズムとして、「Laplace メカニズム」と呼ばれる方式が知られている。Laplace メカニズムの分割表への適用は極めて容易であり、具体的には分割表における各セル (値が 0 のセルを含む) に対し、0 を中心とした Laplace 分布[8]に従う乱数 (Laplace ノイズ) をそれぞれ独立に加えれば良い。

しかし、国勢調査の統計表などの大規模な分割表に対して単に Laplace メカニズムを適用すると、以下のような実用上の課題が生じ、統計の有用性が損なわれることが知られている (寺田他(2015), 伊藤・寺田(2020))。

1. Laplace メカニズムが適用された出力は (実際の人口データではありえない) 負数を多く含みうること (非負制約の逸脱)
2. 非構造的ゼロであったセルのほぼすべてが 0 以外の値を持つようになることから、国勢調査のような大規模なスパースデータに適用すると出力のデータ量が著しく増大すること (スパース性の喪失)
3. 個々のセルに対して一律にノイズが加算されるため、たとえば「ある地域における人口の総数」など、複数セルの部分和を取った際の誤差が大きくなり、精度が劣化すること (部分和精度の劣化)

上記の課題のうち、1 点目の非負制約の逸脱に対しては、単なる負値の 0 への切り上げにより容易に解決可能に思われる。この操作により、Laplace メカニズムの適用結果として負値をとっていたセルの値がすべて 0 に補正されることから、2 点目の課題も緩和される。その一方で、このような単純な「負値の切り上げ」による解決は、各セルの値に対して少しずつの過大バイアスをもたらす。その影響はそれぞれのセル単体での誤差としては小さくとも、それらを足し合わせた小計である部分和、たとえば基本単位区ごとの人口統計における市区

[8] 両側指数分布とも呼ばれる。なお、分布のスケールは、プライバシー損失予算 ε の値と、「大域敏感度 (GS: global sensitivity)」と呼ばれる、問合せの種類によって定まる値に従う。

町村人口などに大きな誤差を生じさせることに繋がりうる。したがって、3 点目の課題を解決することなく、かえって悪化させることが懸念される。

また、伝統的な統計的開示制御手法の一つである PRAM (post randomization) は差分プライバシーを満たすことが知られており、さらに Laplace メカニズムと異なり個票データのレベルで処理を実施することから、その出力となる集計表が非負制約を逸脱することはない。しかし、PRAM はプライバシー効率が極めて悪く、Laplace メカニズム等に対して同じ安全性条件においてより誤差が著しく大きくなるという問題が指摘されている (寺田他 (2017a))。

3.2 ウェーブレット変換に基づく手法

これらの問題を解決するアプローチとして、メッシュ人口統計に対しては局所性保存写像と非負精緻化を伴うウェーブレット変換に基づく手法である非負ウェーブレット法 (寺田他 (2015)) が有効であることが示されている。

非負ウェーブレット法により得られた人口データは、ウェーブレット変換の性質から部分和の精度を制御可能であるとともに、非負精緻化の過程でデータの非負性とスパース性も復元されるという特徴を持つことから、前述の 3 つの課題を同時に解決することが期待される。実際に、2010 年の国勢調査におけるメッシュ人口統計データへの適用に関する実証実験 (伊藤・寺田(2020)) により、非負ウェーブレット法が上記で述べた 3 つの課題をいずれも解決することが示されている。また、Ito et al. (2020) によれば、メッシュ統計においては制約つき最適化に基づくトップダウン構成法 (後述) より高い有用性を備えることが示されている。

しかし、その一方で、非負ウェーブレット法のメッシュ統計以外の統計表への適用方法は自明ではない。非負ウェーブレット法は対象となる分割表のデータに対してウェーブレット変換が適用可能であることを前提としている。標準地域メッシュに従うメッシュ統計の場合は、地理空間上で格子状にデータが配置されていることからウェーブレット変換の適用は容易であるが、たとえば市区町村や町・字、基本単位区などごとに集計された小地域統計に対して非負ウェーブレット法をそのまま適用することは困難である。

3.3 制約つき最適化に基づく手法

前節で示した 3 つの課題を解決する、メッシュ統計以外に対しても適用可能なアプローチとして、制約つき最適化を用いる方法が挙げられる。これは、Laplace メカニズムなどによるノイズの付与後に総数制約[9]と非負制約を制約条件とする最適化を実施し、その解を出力とするものである。

Lee et al. (2015) は、数値最適化法の一種である ADMM (alternating direction method of multipliers) を用いてこれを実現するアルゴリズムを与えている。ただし、この手法は大きな計算コストを必要とし、たとえば 4,096 セルのデータへの適用に 21 秒を要した[10]とされる。また、2020 年センサスでは商用ソルバの Gurobi を用いて制約つき最適化を実装しており、その運用にあたっては商用クラウド基盤である AWS (Amazon Web Service) が提供する分散計算システム EMR (Elastic Map-Reduce) を利用し、最大で 100 台の高性能マシン (それぞれが 96 個の仮想 CPU と 768GB の RAM を備える) を用いた大規模なシステムが構築されている[11]。

---

[9] 入力データにおけるレコードの総数 n が、出力される分割表においても保存される (セルの総和が n になる) という制約条件。

[10] CPU に AMD Opteron 2216 (2.4GHz, 2 core)を用いた計測結果 (Lee et al. (2015))。

[11] DAS EMR Configuration (edited on Sep 16, 2021), https://github.com/uscensusbureau/DAS_2020_Redistricting_Production_Code/wiki/DAS-EMR-Configuration (2023 年 3 月 12 日最終アクセス) による。

寺田他 (2017b) は、総数制約と非負制約を制約条件とした最適化問題が、多次元ベクトル空間における正規単体 (canonical simplex) への射影問題に帰着される[12]ことを利用して、大規模なデータに対して効率的にこれらの制約を満たす差分プライバシーの適用手法を示している。この手法に基づき、当時の一般的なノート PC を用いて 100,000 セルのデータへの適用を 12.6 ミリ秒で完了する[13]実装方法が報告されており(寺田他 (2017a))、国勢調査などの大規模な統計に適した手法と言える。

これらの制約つき最適化を国勢調査の統計表に適用するにあたり、一般には最小集計区分(国勢調査の場合は基本単位区) 単位の人口に対してこの方法を適用し、市区町村単位や都道府県単位の人口はボトムアップ的にそれらを畳み上げて集計することにより得る構成法が考えられる。これをボトムアップ構成法と呼ぶ。また、それとは逆に、まず全国の人口総数を総数制約とした都道府県単位の人口に対してこの方法を適用してプライバシー保護済みの都道府県単位の人口を算出し、次にその (プライバシー保護済みの) 都道府県単位の人口を総数制約としてプライバシー保護済みの市区町村単位の人口を算出するようなやり方で、トップダウンの方向で再帰的に適用する構成法も考えられる。

本稿では、前者をボトムアップ構成法、後者をトップダウン構成法と呼ぶ。なお、2020 年アメリカセンサスで用いられた TDA (top down algorithm) はトップダウン構成法によるアルゴリズムの一種と言える[14]。ボトムアップ構成法とトップダウン構成法のいずれも、制約つき最適化を用いることにより非負制約は充足され、また最適化により非負制約を充足する過程でスパース性も回復することが期待される。さらに、トップダウン構成法は部分和精度の劣化を併せて解決することが期待される[15]。

いずれの構成法も、その適用対象はメッシュ統計に限定されておらず、それ以外の統計にも適用可能である。ただし、メッシュ統計以外についてわが国の国勢調査に適用した実証研究はなされておらず、その実用性や定量的な性質は明らかではない。

4. 2015 年の国勢調査データへの差分プライバシーの適用に関する実証実験

前節での議論の通り、わが国の国勢調査に適した差分プライバシーの適用方式について、メッシュ統計についてはウェーブレット変換に基づく手法の有用性が定量的に実証されているが、それ以外の統計への適用についてはどのような方式が適しているか明らかではない。そこで、本研究では 2015 年国勢調査における個票データを用いて差分プライバシーの適用に関する比較実験を実施した。本節では、本実験で用いた対象データと実験手順を示す。

4.1 実験データ

本実験では、2015 年の国勢調査における個票データ (全数データ)から、集計区分が異なる 3 種類の小地域集計表を作成し、これを実験における対象データとした。それぞれの集計表は、いずれも地域に関する集計区分を基本単位区とした人口に関する集計表であり、属性に

---

[12] この最適化問題は、入力データのレコード総数を n、ノイズ付与後のすべてのセル値を並べたベクトルを $\boldsymbol{w}\ (\in \mathbb{R}^p)$ としたとき、$\arg\min_{\boldsymbol{\varphi}} \|\boldsymbol{w}-\boldsymbol{\varphi}\|_2^2\ s.t. \boldsymbol{\varphi} \geqslant 0, \|\boldsymbol{\varphi}\|_1 = n$ として表される。この式において、制約条件部分は p 次元ベクトル空間上の正規単体 Δ^p をn倍に拡大した単体を意味することから、この問題は正規単体 Δ^p 上における $\boldsymbol{w}/n$ との最近傍点、すなわち $\boldsymbol{w}/n$ から Δ^p への射影を求める問題（正規単体への射影問題）に帰着できる。

[13] CPU に intel Core i5-6200U (2.3GHz, 2 core)を用いた計測結果。

[14] 厳密には、TDA は少数民族 (AIAN: American Indian and Alaska Native) や、各種の法的な要請による不変値 (invariants) に対応するための特殊処理なども含む。

[15] ただし、その代償としてボトムアップ構成法より強いノイズを必要とする。

関する集計区分が異なる (人口のみ、男女別、男女・年齢 5 歳階級別の 3 種類)。具体的には、①基本単位区別人口、②基本単位区・男女別人口、③基本単位区・男女・年齢 5 歳階級別人口の 3 種類の集計表を作成し、これらを対象データとして実験を行う。

以降において、基本単位区別人口を集計表 1、基本単位区・男女別人口を集計表 2、基本単位区・男女・年齢 5 歳階級別人口を集計表 3 とそれぞれ呼ぶことにする。

4.2 実験の方法

本実証では、集計表 1～集計表 3 を対象として各種の差分プライバシーの実現手法を適用し、その有用性を比較評価した。以下、適用にあたっての具体的な方法と比較評価にあたっての評価指標を説明する。

まず、適用対象とする差分プライバシーの実現手法として、PRAM、Laplace メカニズム、トップダウン構成法、ボトムアップ構成法の 4 種類とした。ウェーブレット変換に基づく手法は、メッシュ統計には有効な手法であるが、それ以外の地域区分を持つ集計表への適用は困難であるため、対象からは除いている。なお、前述したように Laplace メカニズムの出力は「負の人口」を含みうる (非負制約を満たさない) ため、負の人口を 0 に切り上げる事後処理を併せて施している。トップダウン構成法やボトムアップ構成法における制約最適化手法としては、正規単体への射影に基づく手法 (寺田他 (2017b)) を採用した。ここで、集計表 2 および集計表 3 へのトップダウン構成法の適用にあたっては、属性に関する集計区分ごとの適用結果を併合することにより作成した。たとえば集計表 2 へのトップダウン構成法の適用にあたっては、男性の人口に対してトップダウン構成法を適用した結果と、女性の人口に対してトップダウン構成法を適用した結果を一つの表として併せたものを、集計表 2 への適用結果としている。

また、これらの手法の適用にあたってのプライバシー損失予算 (ε) は、0.1, 0.2, 0.7, 1.0, 1.1, 5, 10, 20 の 8 種類を設定し、それぞれの値のもとで前述の 4 種類の手法を適用した。ここで、0.7 と 1.1 は、それぞれ (プライバシー損失予算の値として議論されることが多い) $\log_e 2$, $\log_e 3$ の近似値として採用したものである。なお、PRAM とトップダウン構成法は、地域区分ないし属性区分ごとに異なるプライバシー損失予算を配分するよう構成することもできるが、本実験では均等に配分するものとした。

有用性の評価にあたっては、実際の集計データの活用の実態に鑑み、基本単位区ごとの (最も細かい) 人口だけでなく、より大きな地域区分ごとの人口 (部分和) に関する有用性も併せて評価した。具体的には、都道府県ごとの人口、市区町村ごとの人口、町字ごとの人口、基本単位区ごとの人口について、その誤差を定量的に比較する。誤差指標としては平均絶対誤差 (MAE, Mean Absolute Error)と二乗平均平方根誤差 (RMSE, Root Mean Squared Error) を用い、集計表の種類 (3 種類)、適用する手法 (4 種類)、プライバシー損失予算 (8 種類)、部分和をとる地域区分 (4 種類) ごとにそれぞれ計算した。

なお、実験で利用した計算機環境の制約から、本実験では (日本全国ではなく) 都道府県を最上位の地域区分とした。すなわち、日本全国に関する集計表を作成して各手法を適用するのではなく、都道府県ごとの集計表を 47 個作成し、それらに対して独立に各手法を適用した上で、最後にそれらの結果を統合して上記の評価指標を計算している[16]。

[16] たとえば、本来のトップダウン構成法では、都道府県よりさらに上位の区分として「全国」を設定すべきであるが、本実験では「都道府県」を最上位とし、都道府県→市区町村→町字→基本単位区の 4 段階としている。

5. 実験の結果と考察

表1〜3にMAEを指標とした本実験の評価結果を示す。各表において、(a) PRAM, (b) Laplace, (c) BottomUp, (d) TopDown は、それぞれ PRAM、Laplace メカニズム (+ 負値の切り上げ)、ボトムアップ構成法、トップダウン構成法を指す。なお、本実験においては MAE を指標とした場合と RMSE を指標とした場合で大きな傾向差が見られなかったため、RMSE を指標とした評価結果は割愛する。なお、表 1 において、Laplace メカニズム以外の 3 つの手法における都道府県ごと人口の誤差が 0 となっている点については留意が必要である。これは今回の実験条件による影響であり実務上の意味を持たない[17]ことから、以降の考察においては考慮しないものとする。

第 2 節で述べた通り、差分プライバシーは様々なプライバシー保護手法に対して統一的な安全性保証を与える枠組みであり、実現手法がいかに異なるものであっても、プライバシー損失予算の値が同じであれば同一の安全性が保証されるという性質を持つ。つまり、集計表の種類、プライバシー損失予算、部分和をとる地域区分を揃えた上で、手法ごとの指標を比べると、同一の集計条件と安全性の元で有用性が手法ごとにどの程度変化するかを定量的に比較することができる。

この観点で PRAM を他の手法と比較すると、PRAM は他の手法に対してほとんどの場合においてプライバシー保護の効率が大きく劣ることが確認できる。なお、プライバシー損失予算が小さい場合における、基本単位区ごとの PRAM の適用結果については他より優れているようにも見えるが、これは PRAM の有用性を示すものではなく、セル単位での誤差の評価指標としての妥当性に関する事象である。

たとえば表 3 において、基本単位区ごとの値に関して PRAM が与える誤差は、プライバシー損失予算の多寡に関わらず (ε=0.1〜20 のすべてにわたって) ほぼ変化しない。これは、プライバシー損失予算をその数値に抑えるために必要な PRAM による攪乱の程度が極めて大きく、ほぼ完全にランダムな一様分布に従って人口が分布するような出力となっていることを反映している。

PRAM における攪乱は、個票データのある属性値に関し、ある定められた確率 (維持確率) ρ でその値をそのまま維持し、確率 1-ρ でその値をランダムに置き換える。ここで、PRAM が差分プライバシーを満たすとき、ρ の値は ρ と ε の関係式によって定まるが、実際にこの値を計算すると、ε の値をよほど大きくしない限り、元の値が維持される確率 ρ の値はほぼ 0 に近い値になることが多い。このとき、ε の値を多少大きくして安全性を緩和してもその出力の性質はほとんど変化せず、ランダムな一様分布なままに留まる。

ここで、表 3 における基本単位区ごとの集計表は、その値のほとんどが 0 もしくは 1 となるような、かなりスパースなものであるため、出力がランダムな一様分布に従うとしても、基本単位区ごとには一見して悪くない精度を持つように見える。しかし、これは単に見せかけの精度であり、実際には統計としてまったく意味をなさない。たとえば、同じく表 3 において、町字ごと、市区町村ごとなどの地域区分における部分和に関して PRAM が与える誤差

[17] これらの 3 つの手法 (PRAM、ボトムアップ構成法、トップダウン構成法) は、いずれも入力となるレコードの総数を保存する性質を持つ (総数制約の充足)。第 4.2 節で述べた通り、本実験では (環境の制約から) 都道府県を最上位の地域区分としている。そのため、(属性に関する集計区分を持たない) 集計表 1 に対して、これらの総数制約を充足する手法を適用すると、総数制約の充足により最上位の地域区分の人口の誤差は 0 となる。すなわち、表1 において都道府県ごとの誤差が 0 となるのは今回の実験条件がもたらした結果であり、本来の通りに日本全国を最上位の地域区分として集計する場合には、代わりに日本全国の総人口の誤差が 0 となる。

表 1　集計表 1 (基本単位区別人口) に対する評価結果 (MAE)

ε	手法	MAE(都道府県)	MAE(市区町村)	MAE(町字)	MAE(基本単位区)
0.1	(a)PRAM	0.00	14408.44	520.10	48.65
	(b)Laplace	98607.22	2490.96	83.50	17.60
	(c)BottomUp	0.00	855.53	72.06	17.38
	(d)TopDown	0.00	79.09	73.35	49.83
0.2	(a)PRAM	0.00	14399.53	520.26	48.64
	(b)Laplace	30844.00	817.25	39.15	9.23
	(c)BottomUp	0.00	367.38	36.86	9.21
	(d)TopDown	0.00	41.12	37.78	30.42
0.7	(a)PRAM	0.00	14401.57	520.09	48.65
	(b)Laplace	5433.01	157.62	11.06	2.75
	(c)BottomUp	0.00	97.37	10.81	2.75
	(d)TopDown	0.00	11.62	11.26	10.42
1	(a)PRAM	0.00	14406.80	520.17	48.65
	(b)Laplace	3483.00	104.64	7.65	1.92
	(c)BottomUp	0.00	65.78	7.52	1.91
	(d)TopDown	0.00	7.84	7.83	7.38
1.1	(a)PRAM	0.00	14400.56	520.19	48.64
	(b)Laplace	3117.37	94.48	6.96	1.74
	(c)BottomUp	0.00	57.97	6.84	1.74
	(d)TopDown	0.00	7.13	7.12	6.73
5	(a)PRAM	0.00	14351.58	518.16	48.47
	(b)Laplace	609.34	19.08	1.54	0.39
	(c)BottomUp	0.00	13.10	1.51	0.38
	(d)TopDown	0.00	1.60	1.57	1.52
10	(a)PRAM	0.00	10215.78	359.65	34.59
	(b)Laplace	314.63	9.65	0.77	0.19
	(c)BottomUp	0.00	6.43	0.76	0.19
	(d)TopDown	0.00	0.84	0.79	0.76
20	(a)PRAM	0.00	3.53	0.23	0.02
	(b)Laplace	152.12	4.80	0.38	0.10
	(c)BottomUp	0.00	3.15	0.38	0.10
	(d)TopDown	0.00	0.42	0.39	0.38

表 2　集計表 2 (基本単位区・男女別人口) に対する評価結果 (MAE)

ε	手法	MAE(都道府県)	MAE(市区町村)	MAE(町字)	MAE(基本単位区)
0.1	(a)PRAM	35301.70	7277.54	261.95	24.80
	(b)Laplace	158482.08	3940.62	96.57	16.10
	(c)BottomUp	4800.07	977.40	68.02	15.55
	(d)TopDown	60.08	79.80	69.48	36.40
0.2	(a)PRAM	34432.49	7273.24	261.97	24.81
	(b)Laplace	50430.20	1271.48	41.92	8.77
	(c)BottomUp	2081.00	443.80	36.11	8.68
	(d)TopDown	24.09	39.48	36.56	24.83
0.7	(a)PRAM	29972.11	7260.04	261.72	24.79
	(b)Laplace	7016.17	193.09	11.17	2.72
	(c)BottomUp	432.39	100.55	10.83	2.71
	(d)TopDown	9.27	11.83	11.16	9.67
1	(a)PRAM	27463.17	7255.24	261.68	24.80
	(b)Laplace	4239.02	120.71	7.69	1.90
	(c)BottomUp	310.38	68.29	7.50	1.89
	(d)TopDown	7.14	8.10	7.77	7.00
1.1	(a)PRAM	26432.21	7247.65	261.63	24.80
	(b)Laplace	3728.69	107.08	7.00	1.73
	(c)BottomUp	271.13	61.45	6.84	1.73
	(d)TopDown	5.63	7.30	7.07	6.43
5	(a)PRAM	5492.60	7215.53	261.27	24.78
	(b)Laplace	653.28	19.92	1.54	0.38
	(c)BottomUp	59.69	12.58	1.51	0.38
	(d)TopDown	1.19	1.58	1.57	1.51
10	(a)PRAM	530.02	7197.73	260.39	24.70
	(b)Laplace	315.57	9.95	0.77	0.19
	(c)BottomUp	26.27	6.56	0.76	0.19
	(d)TopDown	0.54	0.82	0.78	0.76
20	(a)PRAM	7.51	5115.43	180.82	17.74
	(b)Laplace	159.90	4.92	0.39	0.10
	(c)BottomUp	14.20	3.23	0.38	0.10
	(d)TopDown	0.25	0.40	0.39	0.38

表 3　集計表 3 (基本単位区・男女・年齢 5 歳階級別人口) に対する評価結果 (MAE)

ε	手法	MAE(都道府県)	MAE(市区町村)	MAE(町字)	MAE(基本単位区)
0.1	(a)PRAM	15899.43	587.27	18.50	2.05
	(b)Laplace	376314.96	9323.57	173.38	10.77
	(c)BottomUp	14008.77	544.73	25.62	3.23
	(d)TopDown	81.64	76.50	30.75	3.44
0.2	(a)PRAM	15874.93	586.70	18.49	2.05
	(b)Laplace	175973.88	4359.93	81.69	5.69
	(c)BottomUp	11884.21	447.52	19.48	2.80
	(d)TopDown	41.25	39.09	20.72	3.28
0.7	(a)PRAM	15703.82	584.13	18.46	2.05
	(b)Laplace	40261.21	997.68	19.59	1.90
	(c)BottomUp	5618.71	203.17	8.85	1.55
	(d)TopDown	11.51	11.42	8.38	2.66
1	(a)PRAM	15573.72	582.26	18.44	2.05
	(b)Laplace	25349.47	628.32	12.71	1.38
	(c)BottomUp	4077.82	146.72	6.56	1.20
	(d)TopDown	7.94	7.93	6.20	2.37
1.1	(a)PRAM	15540.75	581.12	18.43	2.05
	(b)Laplace	22484.80	557.35	11.37	1.27
	(c)BottomUp	3725.80	133.89	6.05	1.12
	(d)TopDown	7.26	7.28	5.73	2.29
5	(a)PRAM	12827.43	541.94	17.99	2.04
	(b)Laplace	3506.22	87.20	2.10	0.31
	(c)BottomUp	795.49	28.69	1.50	0.30
	(d)TopDown	1.61	1.60	1.45	0.96
10	(a)PRAM	6345.74	469.39	17.20	2.02
	(b)Laplace	1684.23	41.92	1.03	0.16
	(c)BottomUp	395.63	14.25	0.76	0.15
	(d)TopDown	0.77	0.80	0.74	0.55
20	(a)PRAM	356.10	433.32	16.58	1.98
	(b)Laplace	839.77	20.90	0.52	0.08
	(c)BottomUp	197.79	7.14	0.38	0.08
	(d)TopDown	0.40	0.40	0.38	0.29

を見ると、誤差の累積により精度が大きく劣化し、元の集計表の性質が大きく損なわれていることが読み取れる。

また、Laplace メカニズムの評価結果は、最小の集計区分である基本単位区ごとの人口の誤差の観点では、ボトムアップ構成法とほぼ並んで優れた結果を得ている。しかし、それらを畳みあげた部分和である、町字ごと、市区町村ごとなどの上位の地域区分における人口については誤差が顕著に拡大する。これは、第 3.1 節で議論した通り、非負制約を充足させるための負値の切り上げが、結果数値の全体に「広く薄く」正のバイアスを発生させることを反映している。このバイアスは、基本単位区レベルの結果数値に対してはあまり大きな誤差として見えることはないが、より大きな地域区分に対しては、致命的なレベルでの過大推計をもたらしうることが確認できる。

ボトムアップ構成法は、Laplace メカニズムと似た傾向を持つが、町字ごとや市区町村ごとなど、地域区分が大きくなるにつれて、Laplace メカニズムと比較して誤差が大きく改善される。これは上記の議論の裏返しで、制約つき最適化の効果によりバイアスの発生が抑制された効果によるものと考えられる。

トップダウン構成法の評価結果の傾向は上記のいずれとも異なる。基本単位区で見た場合の誤差はボトムアップ構成法に劣るものの、上位の地域区分における結果数値に誤差が累積することなく、地域区分のレベルにかかわらずほぼ同一の誤差に抑えられており、結果数値の有用性はより高まっていると言える。たとえば、2020 年アメリカセンサスで用いられたものと類似したプライバシー損失予算（ε=20）における実験結果では、地域に関する単変量および変数の組み合わせで設定される結果数値のいずれにおいても、あらゆる地理的区分の部分和における誤差は 1 以下となっており、有用性が相対的に高くなる結果が得られている。

これらの結果から、ボトムアップ構成法とトップダウン構成法のいずれであっても、制約つき最適化に基づく手法は国勢調査への適用における有用性の向上に一定の効果を持つことが確認できた。いずれの構成法が優れているかについては、どの集計区分における精度を重視するかによる。すなわち、基本単位区レベルでの結果数値の誤差から見た場合には、ボトムアップ構成法が相対的に高い有用性を持つと言えるが、市区町村や町字などの、より大きな集計区分における誤差も一定に抑えたい場合は、トップダウン構成法がより優れた性質を備えると言える。

なお、これらの各手法に対する比較評価の議論からわかるように、差分プライバシーの実現手法の有用性に関して議論する上では、セル単位での誤差 (MAE や RMSE など) だけを比較するだけでは不十分であると言える。たとえば PRAM の評価結果が示すように、集計単位が「細かすぎる」場合には、出力が単なる一様ランダムに従うような意味がないものであっても、セル単位で誤差を評価すると良好な結果を得ているように見えてしまうことがある。また、Laplace メカニズムの評価結果に関する議論で示した通り、出力に「広く薄く」バイアスが乗せられていても、セル単位の誤差の値にはその影響が明確には反映されない。

したがって、差分プライバシーの実現方式の比較検証にあたっては、セル単位での誤差を比較評価するだけでなく、たとえば本実験において部分和誤差を検証したように、異なる観点からの統計量についても併せて検証することが重要となりうる。この点について、次節でさらなる議論を加える。

6. 差分プライバシーの有効性の評価方法に関する試論的考察―生態学的誤謬を例に―

トップダウン構成法を含む各種の差分プライバシーの実現方式を適用することによって作成されたノイズ付加済みのデータの有用性については、元データからの誤差の程度を把握

するために、第 4 節と第 5 節で議論したように、平均絶対誤差(MAE)や 2 乗平均平方根誤差(RMSE)などのセル単位での平均誤差を用いて、定量的な評価が行われることが少なくない。それに対して、第 5 節で議論したように、基本単位区といった粒度が細かな地域区分の場合、元データに含まれる統計数値と差分プライバシーの実現方式が適用された数値とのずれをセル単位での平均誤差だけでは適切に評価することができないことが指摘される。したがって、国勢調査への差分プライバシーの適用可能性を追究するにあたっては、セル単位の MAE や RMSE だけでなく、それとは異なる統計量を用いた有用性の検証を行うことが求められる。

その場合の有用性評価に関する 1 つの考え方は、異なる粒度の地域区分で集計されたデータに含まれる統計数値のおのおのに攪乱を施した場合に、各データのそれぞれの分布特性が元データの特性からどの程度乖離しているかを比較・検証し、その乖離が小さいほど集計データに含まれる統計数値の有用性が高いと判断することである。本稿における国勢調査を用いた差分プライバシーの実現方式に関する比較研究では、プライバシー損失予算εの大きさだけでなく、地域区分の粒度によって、差分プライバシーに基づいて攪乱された統計数値と元データにおける統計数値との誤差に関する分布の傾向が異なることが確認された。ゆえに、差分プライバシーに伴う攪乱によって統計数値の安全性が保証された上で、様々な地域区分において、集計されたデータと個票データとの分布特性の差異の程度を把握することによって、地域区分を考慮した場合の有用性の評価が可能になる。具体的には、地域区分の粒度を変えた場合に、個票データに基づく変数間の相関性といった有用性の指標と集計データから算出された場合の指標を比較・検討することが考えられる。それによって、差分プライバシーが適用されたノイズ付与済みの集計データの作成の際に用いられた各種の地域区分の中で、元の個票データと比較して最も近似的なデータにおける地域区分が選定される。

このような地域区分の設定に関しては、集計データと個票データとの関係性が注目される。この関係性は、2 つの社会的集団間に見られる因果関係が、マクロレベルとミクロレベルにおける次元の相違として位置付けられる生態学的誤謬(ecological fallacy)として概念化されてきた(伊藤(2002))。Robinson によって提唱された生態学的誤謬とは、個別的主体群の社会経済的特性に関する「個別的相関(individual correlation)」、および地域レベルの集団的特性について算出された「生態学的相関(ecological correlation)」の論理的な次元の違いに起因する形で、個別的主体の社会経済的属性間の関連性を把握するために、生態学的相関によって個別的な行為特性を把握することである(Robinson(1950), 伊藤(2002), 伊藤(2011))。

Cohen et al.(2022)では、生態学的誤謬論から展開された生態学的回帰(ecological regression)や生態学的推論(ecological inference)の観点から、差分プライバシーの実現方式である TDA の有効性を検証していることが興味深いと言える。生態学的回帰は、Goodman(1953)によって提唱された手法であり、生態学的変数間における回帰分析を行うことによって、個別的主体の社会的な行為事象に関する傾向的な把握を試みる手法である(伊藤(2011, p.40))[18]。さらに、生態学的推論とは、「集計のレベルが生態学的分析(ecological analysis)に及ぼす集計効果を把握し、その集計効果を調整することによって、集計データを用いても個々人の社会経済的属性

[18]「集計データの利用の背景にあるのは、「生態学的誤謬の慎重な適用」(Allardt(1969, p.42))という観点から、「集計効果(aggregation effects)(Holt et al.(1996))」の影響を考慮することによって、個々人の社会的行為に対して集計データによる実証的な社会研究の可能性を追究しているということである。・・・(中略)・・・「生態学的誤謬の慎重な適用」という視点に立てば，「生態学的データや生態学的相関の分析は，個別的行為に関する叙述を行う上で強力な道具になる」だけでなく、「それに対応する個票データ(individual data)よりも実りのある因果的解釈を容易に」しうる」(伊藤(2011，p.40))。

と社会的行為との関係を推測する」ことである(伊藤(2011, p.41))[19]。

なお、Cohen et al. (2022)において検証を行っている生態学的回帰は、Goodman(1953)の議論に基づいて、以下のとおり説明することができる。すなわち、生態学的回帰の手法は、一般に個別的主体の社会的な行為事象に関連した研究の代用としては利用されないが、ある一定の条件のもとで、個別的主体の行為に関する推論を行うために、「生態学的変数」間における回帰分析の成果を利用することは可能であると結論づけている(Goodman(1953, p.664))。したがって、生態学的回帰では、ある一定の条件のもとでは、生態学的変数間における回帰分析を行うことによって、個別的主体の行為に関連した推論が行われうる(Goodman(1953, p.664), 伊藤(2011))。

なお、Goodman(1953)による生態学的回帰の説明においては、最初に人種と識字状況に関する 4 分割表(図 1)に基づいて、つぎの式(1)が設定されている。

$Y = r + (p - r)X$・・・(1)

ここで

Y=$(a + b)/T$・・・各地域における非識字率

X=$(a + c)/T$・・・各地域における黒人比率

p=$a/(a + c)$・・・黒人のなかで非識字である者の比率

r=$b/(b + d)$・・・白人のなかで非識字である者の比率

つぎに、式(1)において、最小 2 乗法を用いてパラメータ p と r が推定されることによって、同時分布における度数が導出される(伊藤(2011, p.41))。

Cohen et al. (2022)では、異なるパラメータ ε をもとに差分プライバシーの実現方式が適用された、地域区分が異なるセンサスデータを対象に生態学的回帰を行っている。Cohen 等の研究の目的は、差分プライベートなデータに基づく適切な区画改定が可能であるかを検証することである。本研究においては、2010 年センサスに基づいて、アリゾナ州の 2 つの郡(ナバホ郡(Navajo County)とピマ郡(Pima County))で把握された住民を対象に、住民に関する情報(居住地域、民族、性別、年齢、人種)に基づいて生態学的回帰を行っている。なお、ナバホ郡の人口は、107,449 人(白人系が 44%、ヒスパニック系が 11%、ネイティブアメリカン/インディアン系が 42%)、ピマ郡の人口は、980,263 人(白人系が 55%、ヒスパニック系が 35%、ネイティブアメリカン/インディアン系が 2.5%)である。パラメータ ε=2 と ε=19 の場合、ナバホ郡とピマ郡のそれぞれにおいて、ランダムに選挙区(district)を選定した上で、ウェイト[20]を付けた生態学的回帰(weighted ecological regression)が行われている。表 4 は、ナバホ郡とピマ郡において、2020 年の大統領選挙において、選挙区別に見たジョー・バイデン氏への投票率と選挙区におけるネイティブアメリカンあるいはヒスパニックの比率から、ウェイトを付けた生態学的回帰によってナバホ郡とピマ郡における人種別の投票率を推定したものである。なお、生態学的回帰による推定結果は、16 回繰り返して行った結果に基づいている。

表 4 に示されるように、重みをつけた生態学的回帰の結果においては、ノイズがない場合と比較して、ノイズを含む結果については、大きな差異が見られないことが確認できる。ピマ郡の場合、相対的に大きなノイズが付与される ε =2 において得られるノイズありの結果は、ノイズなしの結果と比較して少し差が見られるものの、それは誤差の範囲とみなすこともできる。この検証結果から、Cohen et al.(2022)においては、ウェイトを考慮した生態学的回帰を行うことによって、実務担当者が差分プライベートなデータを用いた場合でも、地域ご

[19] 生態学的推論に関する先行研究としては、例えば、Openshaw や Holt 等による生態学的誤謬論の観点に立った集計効果の定量的な研究がある(Openshaw(1984), Holt et al.(1996))。

[20] Cohen et al.(2022)においては、ウェイトとして地区の人口や有権者数が想定されている。

図 1　識字状況と人種の関係に関するイメージ図

		人種 黒人	白人	
識字状況	非識字	a	b	a+b
	読み書き可能	c	d	c+d
		a+c	b+d	T

注　Goodman(1953)をもとに一部修正

表 4　人種別のジョー・バイデン氏への投票率に関する生態学的回帰の推定結果

地域/人種		ε =2		ε =19	
	ノイズなし	最小値	最大値	最小値	最大値
ナバホ郡					
ネイティブアメリカン	0.886	0.887	0.892	0.885	0.886
非ネイティブアメリカン	0.169	0.167	0.170	0.169	0.169
ピマ郡					
ヒスパニック	0.661	0.653	0.663	0.659	0.662
非ヒスパニック	0.573	0.572	0.575	0.572	0.573

注　Cohen et al.(2022)、表 5 に基づいている。なお、表頭で示される最小値と最大値は、16 回の生態学的回帰の試行によって得られた推定値群における最大値と最小値を示している。

との居住人口における人種別の分布に基づく区画改定の妥当性が検証可能であることが議論されている。

本研究は、個票データを用いた場合のミクロレベルの回帰分析の結果に近似的であるためには、どのようなレベルでアグリゲーションを行い、その集計データに差分プライバシーに基づくノイズをどの程度導入するのがよいかを検証することを目指した実証研究の事例の１つと位置付けることもできる。換言すれば、生態学的誤謬に起因する誤差と差分プライベートなノイズに基づいて適切な地域区分が設定されていれば、ミクロレベルの検証結果に近づくことが期待できることを示唆している。

7. まとめ

本稿では、公的統計における差分プライバシーの方法的な有効性を追究するために、最初に、アメリカを例に、公的統計に対するプライバシー保護に関するこれまでの動向と差分プライバシーの適用における社会的背景を概観し、差分プライバシーの方法論が適用された統計データにおける秘匿性と有用性に関する定量的な評価方法について論じた。また、Christ et al. (2022)を考察の素材として、各種のスワッピング技法と差分プライバシーの実現方式を比較対象とした上で、有用性の指標としては MSE と KL 情報量を、秘匿性の指標についてはリンケージ技法を用いた検証の概要を紹介した。Christ らによれば、差分プライバシーを適用した場合でも、スワッピングが適用されたデータと同様の精度が保持できることが確認されているが、本研究は、センサス局が、2020 年センサスでプライバシー保護の方法をスワッピングから差分プライバシーに変更したことについて、実証結果で支持したものと言える。

つぎに、本稿においては、わが国の公的統計における差分プライバシーの適用可能性を追究するために、国勢調査の個票データを用いて、地理的区分が異なる統計表、各種の差分プライバシーの実現方式を適用した場合の有用性の評価を行った。本研究では、国勢調査に差分プライバシーを適用する場合には、トップダウン構成法の場合、他の手法と比較して、評価指標として MAE を用いた場合の有用性が相対的に高いことが確認できた。このことは、

国勢調査のような階層的な地域区分を有するデータにおいて、最上位の地域区分でノイズを生成し、トップダウンで調整を図りながら統計表の各セルにノイズを割り当てる場合、最小地域区分ごとの集計表のセルにノイズを付与し、畳み上げて集計した場合よりも、相対的に精度の高い数値が得られることを意味している。

本稿では、Robinson の生態学的誤謬論の観点から、差分プライバシーに基づくノイズ付与済みの集計データにおける有用性の評価についても議論を行った。それは、様々な地域区分において、集計されたデータと個票データとの分布特性の差異の程度を捉えた上で、個票データと比較して最も近似的なデータ構造を与える地域区分を有する集計データの利用可能性を追究するものである。Cohen et al.(2022)においては、ウェイトを考慮した生態学的回帰を行うことによって、差分プライベートなデータから、区画改定を行うことが可能であることが確認されている。このことは、集計データに対して差分プライバシーに基づくノイズを適切に付与することができれば、ミクロデータを用いない場合でも、集計データからミクロレベルのより精度の高い分析結果が導かれることを意味している。これについては、さらに個票データを用いて、アグリゲートされたデータの地域区分の相違、ノイズの付与の程度が生態学的推論にどのような影響を及ぼすかについての精密な検証が必要になるだろう。

なお、国勢調査による検証結果で示されたように MAE や RMSE で有用性を評価しようとしても、PRAM で攪乱されたデータの精度が適切に評価できず、PRAM の精度の低さが誤って評価されたことが本研究で確認されている。これについては、MAE や RMSE の評価のあり方についての検討の必要性を示唆している。その一方で、差分プライベートなデータを対象にして、生態学的誤謬論に基づく有用性の評価方法のさらなる検討も含め、別の有用性の評価方法を模索するだけでなく、有用性の評価方法についての体系的な議論も必要なように思われる。これについては、将来的な検討課題としたい。

謝辞

本稿は、2023 年度統計関連学会連合大会における学会報告(2023 年 9 月 5 日)に基づいている。報告内容について貴重なコメントをしていただいた星野伸明先生(金沢大学)に謝意を申し上げたい。また、本稿の一部は、アメリカセンサス局の担当者に対して実施したヒアリング調査(2023年7月20日)をもとに執筆したものである。インタビューに応じていただいた Michael Hawes 氏(アメリカセンサス局)に記して謝辞を表したい。

参考文献

[1] 伊藤伸介(2002),「アメリカにおけるミクロ社会モデルの体系化の試み―オーカットの社会人口モデルと所得移転モデル―」,『統計学』第83号，pp.11-31.

[2] 伊藤伸介(2011),「わが国におけるミクロデータの新たな展開可能性について―イギリスにおける地域分析用ミクロデータを例に―」，明海大学『経済学論集』Vol.23, No.3, pp.36-54.

[3] 伊藤伸介(2022),「ミクロデータの匿名化と統計情報の秘匿可能性について」『経済学論纂(中央大学)』第63巻1・2合併号，pp.1-23.

[4] 伊藤伸介・寺田雅之 (2020), 「詳細な地域データにおける秘匿処理の適用可能性について」『日本統計学会誌』第 50 巻第 1 号, pp.139-166.

[5] 伊藤伸介・寺田雅之・赤塚裕人・北井宏昌(2022),「海外における公的統計に対する攪乱的手法の新たな取り組み―アメリカセンサス局による差分プライバシーの適用を中心に―」『統計研究彙報』第 79 号，pp.131-150.

[6] 伊藤伸介・寺田雅之(2023),「海外における公的統計に関するプライバシー保護の現状―アメリカとイギリスの事例をもとに―」『統計研究彙報』第 80 号, pp.117-136.
[7] 寺田雅之・鈴木亮平・山口高康・本郷節之(2015),「大規模集計データへの差分プライバシの適用」『情報処理学会論文誌』第 56 巻第 9 号, pp.1801-1816.
[8] 寺田雅之・山口高康・本郷節之(2017a),「匿名個票開示への差分プライバシの適用」『情報処理学会論文誌』第 58 巻第 9 号, pp.1483-1500.
[9] 寺田雅之・山口高康・本郷節之(2017b),「高次元大規模データへの差分プライバシー適用のための最適精緻化法」『SCIS2017 予稿集』3B3-5, pp.1-8.
[10] Allardt, E. (1969), "Aggregate Analysis: The Problem of Its Informative Value", Dogan, M. and Rokkan, S.(eds.) *Quantitative Ecological Analysis in the Social Sciences,* the M.I.T Press, London, pp.41-51.
[11] Christ, M., Radway, S., Bellovin, S. M. (2022), "Differential Privacy and Swapping: Examining De-Identification's Impact on Minority Representation and Privacy Preservation in the U.S. Census", Paper Presented at Conference: 2022 IEEE Symposium on Security and Privacy, pp.457-472.
[12] Cohen, A., Duchin, M., Matthews, J.N., and Suwal, B. (2022), "Private Numbers in Public Policy: Census, Differential Privacy, and Redistricting." *Harvard Data Science Review*, Special Issue 2, MIT Press.
[13] Dwork, C. (2007), "An Ad Omnia Approach to Defining and Achieving Private Data Analysis", Proc. 1st intl. conf. Privacy, security, and trust in KDD, pp. 1-13.
[14] Garfinkel, S. Abowd, J. M., and Martindale, C. (2019) "Understanding Database Reconstruction Attack in Public Data", Communications of the ACM, Vol. 62 No. 3, ACM, pp. 46-53.
[15] Goodman, L.A. (1953), "Ecological Regression and Behavior of Individuals", *American Sociological Review*, vol.18, pp.663-664.
[16] Holt, D., Steel, D. G., Tranmer, M., Wrigley, N.(1996), "Aggregation and Ecological Effects in Geographically Based Data", *Geographical Analysis,* Vol.28, No.3, pp.244-261.
[17] Ito, S., Miura, T., Akatsuka, H., and Terada, M. (2020), "Differential Privacy and Its Applicability for Official Statistics in Japan – A Comparative Study Using Small Area Data from the Japanese Population Census", Proc. intl. conf. Privacy in Statistical Databases (PSD 2020), LNCS 12276, Springer, pp. 337–352.
[18] Jamin, R. (2021), "Disclosure Avoidance for the 2020 Census: An Introduction", U.S. Census Bureau.
[19] Lee, J., Wang, Y., and Kifer, D. (2015), "Maximum Likelihood Postprocessing for Differential Privacy under Consistency Constraints." Proc. 21st ACM SIGKDD intl. conf. Knowledge Discovery and Data Mining (KDD '15), pp. 635–644.
[20] McKenna, L. (2019a), "U.S. Census Bureau Reidentification Studies", U.S. Census Bureau.
[21] McKenna, L. (2019b), "Research and Methodology Directorate: Disclosure Avoidance Techniques Used for the 1960 Through 2010 Decennial Censuses of Population and Housing Public Use Microdata Samples", U.S. Census Bureau.
[22] Openshaw, S. (1984), "Ecological Fallacies and the Analysis of Area Census Data", *Environment and Planning A*, Vol.16, pp.17-31.
[23] Robinson, W. S. (1950), "Ecological Correlations and the Behavior of Individuals", *American Sociological Review*, vol.15, pp.351-357.