リサーチペーパー第 58 号　　Research Paper No.58

国勢調査における攪乱的手法の適用とその有効性に関する実証研究

An Empirical Study on the Effectiveness of Perturbative Methods Applied to Japanese Population Census Data

伊藤 伸介
統計研究研修所客員教授
中央大学経済学部教授
ITO Shinsuke
SRTI Guest Professor
Professor, Faculty of Economics, CHUO University

寺田 雅之
統計研究研修所特任教授
(株)NTT ドコモ
TERADA Masayuki
SRTI Specially Appointed Professor
NTT DOCOMO, INC.

加藤 駿典
統計局総務課
滋賀大学大学院データサイエンス研究科修士課程
KATO Shunsuke
General Affairs Division, Statistic Bureau
Master's Course, Graduate School of Data Science, SHIGA University

令和 5 年 4 月
April 2023

総務省統計研究研修所
Statistical Research and Training Institute (SRTI)
Ministry of Internal Affairs and Communications

受理日：令和５年３月 31 日

本ペーパーは、総務省統計研究研修所の客員教授及び特任教授並びに総務省統計局職員である執筆者が、その責任において行った統計研究の成果を取りまとめたものであり、その内容については、総務省統計局又は統計研究研修所の見解を表したものではない。本ペーパーの内容については、執筆者に問い合わせ願いたい。

本研究では、統計法(平成 19 年法律第 53 号)第 32 条の規定に基づき、国勢調査に係る調査票情報を使用した。

国勢調査における攪乱的手法の適用とその有効性に関する実証研究

伊藤 伸介
寺田 雅之
加藤 駿典

概要

公的統計におけるプライバシー保護のための技法として、近年海外では、攪乱的手法(perturbative methods)が積極的に用いられる傾向にある。特に、アメリカセンサス局では、差分プライバシー(differential privacy)の方法論に基づく攪乱の公的統計の実務への適用可能性が追究されてきた。具体的には、2020 年のアメリカ人口センサスにおいて、差分プライバシーを用いた統計表の作成の検討が進められ、再構築攻撃(database reconstruction attack)への対応として、人口センサスを対象に差分プライバシーの方法論の実用性が模索されてきた。

わが国の公的統計に対する差分プライバシーの適用可能性を実証的に追究することは、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有益であると思われる。そこで本稿では、国勢調査を例に、海外での差分プライバシーの検討状況を踏まえつつ、差分プライバシーの方法論のわが国の公的統計データへの有効性に関する比較研究を行った。

特に、平成 27 年国勢調査の個票データを用いて作成した地理的区分が異なる集計表について、各種の差分プライバシーの実現手法を適用した場合の有用性を評価するだけでなく、伝統的な匿名化手法の 1 つである PRAM と比較・検証を行った。その結果から、国勢調査に各種の差分プライバシーの実現手法を適用した場合、トップダウン構成法における有用性が、ボトムアップ構成法と Laplace メカニズムのそれと比較して相対的に高いことが確認できた。したがって、階層的な構造を持つ地理的区分においては、差分プライバシーに基づいて発生するノイズ全体をトップダウンで整合を取りつつ統計表の各セルに割り当てる方法が、相対的に誤差が小さな結果数値をもたらしている。このことは、統計実務の観点から見た場合、差分プライバシーの実現手法としては、トップダウン構成法が有効な手法であることを示唆している。

キーワード：差分プライバシー、アメリカセンサス局、国勢調査、PRAM、有用性、
安全性

An Empirical Study on the Effectiveness of Perturbative Methods Applied to Japanese Population Census Data

ITO Shinsuke
TERADA Masayuki
KATO Shunsuke

Abstract

In numerous countries, perturbative methods are increasingly used as a privacy protection method for official statistics. The U.S. Census Bureau has studied the applicability of perturbation based on differential privacy for official statistics, and empirically investigated the mechanism of differential privacy for the publication of statistical tables created based on the 2020 Population Census data. The U.S. Census Bureau has also examined the applicability of differential privacy for Population Census data as a protection against "database reconstruction attacks".

In discussing future directions for the creation and publication of statistical tables as well as the secondary use of official statistical data, it is important to consider the applicability of differential privacy. Towards this objective, this paper conducts empirical research about the effectiveness of differential privacy for Japanese Population Census data while taking into account the actual situation regarding the application of differential privacy for official statistics in other countries.

Specifically, this research conducts a comparative analysis of data usability values of cells in a statistical table with noise added based on several approaches of differential privacy (with PRAM as a traditional disclosure limitation method) for statistical tables at different geographical levels created using individual data from the 2015 Population Census. The results of this research indicate that when several mechanisms based on the standard of differential privacy are applied to Population Census data, the degree of data usability for the top down method is comparatively better than those for other differential privacy mechanisms such as bottom up method, Laplace mechanism and PRAM. This suggests that adding noise using a top down mechanism generated based on the standard of differential privacy to each individual cell in statistical tables with a hierarchical structure of geographical levels results in cell values with relatively small errors. This in turn implies that from the standpoint of statistical practice, the top down method is more effective than other mechanisms.

Keywords: differential privacy, U.S. Census Bureau, Population Census, PRAM, data usability, security

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技法として、近年海外では、攪乱的手法(perturbative methods)が積極的に用いられる傾向にある(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)では、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の公的統計の実務への適用可能性が追究されてきた。具体的には、センサス局では、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が進められてきた。特に小地域レベルのセンサスの統計表を作成・公表する上で、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差分攻撃(differential attack)[1]への対応として、センサスを対象に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))。

センサス局では、2020 年センサスに関する統計表を公表することで、差分プライバシーの方法論の実用化を果たしたが、ヨーロッパの他の統計作成部局でも、差分プライバシーについての関心が高まっており、差分プライバシーの方法論の可能性が追究されている。欧州統計局(Eurostat)では、差分プライバシーの適用可能性が模索されているが(Bach(2022))、イギリス国家統計局(Office for National Statistics=ONS、以下「ONS」と呼称)やノルウェー統計局といったヨーロッパの統計作成部局では、統計実務の観点から、公表統計表の安全性だけでなく、ミクロデータの提供の可能性について検討を行っている[2]。

海外の動向を踏まえると、差分プライバシーの適用にあたっては、統計表の作成・公表だけでなく、合成データ(synthetic data)のような擬似的なミクロデータの作成において潜在的な可能性が考えられる。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))。

そこで本稿では、国勢調査を例に、海外で注目されている差分プライバシーの方法論を中心に、各種の攪乱的手法を適用した際のデータの特性を実証的に明らかにするだけでなく、わが国の公的統計データに対する攪乱的手法の有用性に関する比較・検証を行うことによって、差分プライバシーの方法論の国勢調査へ適用可能性を模索していきたい。

---

[1] 差分攻撃については, Fraser and Wooton (2005) を参照.

[2] イギリスでは、ONS が死亡データをもとに、センサス局で検討されてきた方法を参考にした上で差分プライバシーの適用可能性に関する検討を行ってきた (Dove(2021))。また、ノルウェー統計局は、プログラム送付型のリモートエグゼキューションに対する差分プライバシーへ適用に高い関心を持っていることが知られている(例えば、Heldal et al.(2019)を参照)。

2.　海外におけるプライバシー保護の動向について

―欧米の例をもとに―

諸外国では、公的統計データに対するプライバシー保護を指向した技術的な対応が多様な様相を呈している。本節では、ヨーロッパとアメリカを例に、公的統計のプライバシー保護をめぐる最近の動きを概括する。

2.1　ヨーロッパにおける公的統計のプライバシー保護の動向

ヨーロッパにおける動向の特徴は、現状としては、差分プライバシーの統計実務への適用に関心をもちつつも、公表される統計表の有用性により配慮をしたことから、差分プライバシーの実用化にまでには至っていないことである(伊藤・寺田(2023))。それに対して、イギリス等では、センサスを対象に、ノイズ付与やスワッピングといった他の攪乱的手法の適用が追究されてきた。公的統計のプライバシー保護の動向としては、以下の 3 点を指摘することができる。

第 1 は、ヨーロッパでは、攻撃者(intruder)のシナリオに基づく特定化の攻撃に対する匿名化措置を施すことによって、匿名化された公的統計ミクロデータの作成・提供が、現在もなされていることである[3]。例えば、イギリスの ONS では、攻撃者が有する外部情報とのマッチングの可能性を想定した上で、ミクロデータに対する匿名化処理がなされる。その場合、特異な属性の組み合わせを有している等、悪意がないような形で個体が偶発的に特定される可能性(偶発的な個体特定(spontaneous recognition))も考慮される[4]。

第 2 は、イギリスでは、2021 年センサスにおいて、集計結果表に含まれるセルに対してノイズを付与する手法である cell key method とミクロデータに対する攪乱的手法の 1 つであるターゲット・スワッピング(targeted data swapping)の実用化を図ってきたことである。例えば、ONS が開発した「オンデマンド型公表システム(Flexible Dissemination System)」は、ウェブ上で利用者が変数群を選択した上で、オンデマンドで集計表を獲得することができるシステムであって、イギリスの 2021 年センサスの多次元集計表の公表方法の 1 つと

---

[3] 以下のウェブサイトを参照されたい。
Policy for Social Survey Microdata
https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/disclosurecontrol/policyforsocialsurveymicrodata

[4] 偶発的な個体特定とは、珍しい属性の組み合わせを持つ個体が、データの利用者によって、偶発的に母集団の中で特定されることである(Duncan et al.(2011))。

して導入されている[5]。このシステムの運用にあたっては、ターゲット・スワッピングが適用されたセンサスの個票データをオンデマンド集計のための元データとした上で、作成された集計結果表に含まれるセルへの cell key method の適用可能性が追究されてきた(Office for National Statistics(2017))。さらにオンデマンド型公表システムには、作成された集計表から個体情報が露見されないような自動化された露見チェック(automated disclosure check)の仕組みを備えることによって、集計表における秘匿の強度を強めている。こうしたチェック済みの集計表を利用者がダウンロードすることができる(伊藤・寺田(2023))。

第 3 には、欧州統計局(Eurostat) を中心に、センサスの統計数値に対する攪乱的手法の可能性が検討されていることである。Eurostat は、差分プライバシーの方法論を中心に、センサスの統計数値に対する攪乱的手法の可能性を検討している。さらに、Eurostat においては、差分プライバシーを適用する上で、秘匿性と有用性の両面から、適用可能なノイズの範囲について検証している(Bach (2022))。

2.2　アメリカにおける人口センサスにおけるプライバシー保護の現状

第 1 節で述べたように、アメリカの場合、センサス局が、「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(differential privacy)の方法論のセンサスへの適用可能性を追究してきた。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(privacy loss budget) ε を設定し、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(TopDown アルゴリズム)に関する検証を進めてきた(Garfinkel et al.(2019), 伊藤・寺田(2020))。これは、後述する「再構築攻撃(reconstruction attack)」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数値の精度も考慮されたことから、センサス局内部に設置された DSEP(=Data Stewardship Executive Policy Committee)での議論等に基づいて、数度の修正がなされた。その後、DSEP が、2020 年センサスの統計表を公表するためにパラメータ ε の最終的な数値に関する決定を行った。具体的には、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)の作成にあたって、「露見回避システム(Disclosure Avoidance System)」で設定されたパラメータの値については、全体のプライ

[5] イギリスの 2021 年センサスにおける多次元集計表の公表スケジュールについては、以下のウェブサイトを参照されたい。
Multivariate data
https://www.ons.gov.uk/census/aboutcensus/censusproducts/multivariatedata

バシー損失予算として ε =19.61 が設定されている(伊藤他(2022))。その後、2021 年 8 月 16 日には 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。

その一方で、アメリカにおけるセンサスのミクロデータの作成状況を概観すると、1960 年センサスにおいて、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、1962 年に最初の一般公開型ミクロデータサンプル(Public Use Microdata Sample=PUMS)として、全人口の 0.1%のレコードが含まれるセンサスのミクロデータの提供が開始された(McKenna(2019))。その後、1970 年から 2010 年までのセンサスにおいては、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だけでなく、ノイズ付与やスワッピングのような攪乱的手法も適用することによって、センサス局は PUMS を作成・公開してきた(伊藤他(2022))。それに対して、差分プライバシーの方法が適用された 2020 年センサスでは、PUMS の作成は未定となっている(2023 年 2 月時点)。なお、センサスの詳細調査票に関する項目は、2000 年センサス以降、American Community Survey(ACS)という形で毎年調査されているが、センサス局は、ACS でも PUMS の作成・提供を行ってきた。さらに、センサス局は，ACS における合成データの作成についての検討を進めていることは興味深い[6]。この合成データにおける差分プライバシーの適用可能性も含め、今後の動向が注目される。

このようにアメリカでは、2010 年センサスまでの「個別的な(ad hoc)」秘匿措置から大きく転換する形で、2020 年センサスでは「フォーマル」なプライバシー保護手法の公的統計の実用性を模索してきたのが、近年の動向と言える。

3. 差分プライバシーの国勢調査への適用

前節で触れたセンサス局における「フォーマル」なプライバシー保護手法の導入に向けた秘匿方針の大きな転換は、オープンデータ化の進展や計算機能力の爆発的な向上、およびこれらを背景としたデータ分析技術の急速な発展から、再構築攻撃を代表とする新たなプライバシー暴露攻撃が現実的な脅威となり、それらへの対策が必要となったことを主な理由としている(伊藤他(2022))。

本節では、センサス局における差分プライバシー導入の理由となった再構築攻撃について、その背景にあるプライバシーリスクの概念である「モザイク効果」と併せて簡単に説明し、それらへの対策としてセンサス局が 2020 年センサスに導入したプライバシーの安全性を評価する枠組みである差分プライバシーについて、その定義と解釈を示す。また、わ

[6] 例えば、以下のウェブサイトを参照。
Disclosure Avoidance Protections for the American Community Survey
https://www.census.gov/newsroom/blogs/random-samplings/2022/12/disclosure-avoidance-protections-acs.html

が国の国勢調査への差分プライバシーの適用を検討する際に生じるであろう技術的課題と、その対策として考えられる実現手法および技術的論点を議論し、それらを踏まえて本研究における研究課題と目標についてまとめる。

3.1　アメリカにおける差分プライバシー導入の技術的背景

モザイク効果 (mosaic effect) とは、「個々のデータセット単体では安全であっても、他のデータセットと組み合わせると（個人の再識別などの）プライバシー暴露を引き起こしうる」という事象を指す言葉である。その存在は潜在的なリスクとして1970年代から知られており(Smith et al. (1996))、アメリカのオープンデータ政策においても、モザイク効果によるプライバシーリスクはデータの公開時に考慮すべき事項として明示的に位置付けられている。たとえばアメリカ行政管理予算局 (Office of Management and Budget) は、2013年5月のオープンデータ政策に関する大統領令とともに公表したメモランダム (Burwell et al. (2013)) において、各政府機関はモザイク効果がもたらすプライバシーリスクを考慮しなければならず、そのために必要な分析は (必要に応じて他の専門機関などの助けを借りつつ) 各機関の責任で実施すべきとした。

また、モザイク効果によるプライバシーリスクを具体的な攻撃手段に適用したものとして再構築攻撃が挙げられる。再構築攻撃は、複数の（それぞれ安全に見える）データを重ね合わせ、そこから導出される制約充足問題を解決することによって、それらのデータに含まれる個人の情報を特定する攻撃手法である。その特徴として、重ね合わせるデータの組み合わせによって暴露される対象や内容が異なり、場合によっては (後述のセンサス局による実験が示す通り) 元となる個票データの大部分が復元されうることが指摘できる。

センサス局は、これらの脅威に鑑み、2010年センサスの集計表に対して再構築攻撃を実験的に適用してその脅威を定量的に分析し、これまでの (スワッピングなどの) 慣用的な秘匿措置では、再構築攻撃の脅威に対してプライバシーを十分に保護できなくなったと結論づけた。以下にその結果の一部を示す (Abowd (2021))。

- 2010年センサスの集計表への再構築攻撃の適用により、アメリカ国民の46%（約1.44億人）の居住ブロック、性別、年代、人種、民族が復元された（年齢に1歳の誤差を許すと71%が復元された)。
- その復元結果を一般に入手可能な市販データと照合することにより、約5,200万人分のレコードについて再識別（個人特定）された。これは、アメリカ国民の約17%に相当する。

再構築攻撃による脅威は、既知の攻撃に対しては安全なデータであっても「想定外」の新しい攻撃に対しては安全と言えないことを示している。つまり、再構築攻撃においては、

重ね合わせるデータの組み合わせが変われば導出される制約充足問題も変わるため、ある特定のデータの組み合わせが再構築攻撃に対して安全だったとしても、別のデータとの重ね合わせに対しても安全とは限らない。つまり、既知のデータだけを考慮しても再構築攻撃によるプライバシー暴露を防ぐことはできない。

これは、少なくとも再構築攻撃によるプライバシー暴露のリスクを考慮すると、「既知の攻撃に対して安全である」と示すだけでは不十分であり、未知の攻撃も含めた安全性を考慮する必要があることを意味している。そのために「個別的な」秘匿措置をいくら重ねても公表する統計の品質を悪化させるだけであり、フォーマルな保証に基づく「包括的な」プライバシー保護の枠組みを導入することが必要となる。

3.2　差分プライバシーの定義と解釈

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性指標を定量的に与える[7]。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。

あるプライバシー保護手法$\mathcal{M}$のプライバシー損失がε以下であることが保証されるとき、$\mathcal{M}$は ε-差分プライバシーを満たすと呼び、より厳密には以下の通り定義される。

定義 1　任意の隣接したデータベース D_1 と D_2 ($D_1, D_2 \in \mathcal{D}$) に対し、ランダム化関数 $\mathcal{M}: \mathcal{D} \to \mathcal{R}$ が下式を満たすとき、$\mathcal{M}$ は ε-差分プライバシー (ε-differential privacy) を満たす。ただし、ここで S は $\mathcal{M}$ の出力空間 $\mathcal{R}$ の任意の部分空間である ($S \subseteq \mathcal{R}$)。

$$Pr[\mathcal{M}(D_1) \in S] \leq e^{\varepsilon} \cdot Pr[\mathcal{M}(D_2) \in S].$$

この定義は、直感的には「A さんのデータが含まれるデータベース D_1 への $\mathcal{M}$ の適用結果と、A さんのデータが含まれないデータベース D_2 への $\mathcal{M}$ の適用結果との見分けがつかなければ、$\mathcal{M}$ の出力が A さんのプライバシーを侵すことはない」と解釈でき、ε が小さいほど「見分けがつきにくい」、つまりプライバシー保護手法 $\mathcal{M}$ が安全であることを意味する (寺田 (2019))。逆に言えば、ε は「$\mathcal{M}$の出力によりプライバシーがどれだけ損なわれるか」を示す指標であるとも言える。この性質から、ε は「プライバシー損失」または「プライバシー損失予算」とも呼ばれる。

定義 1 が示す通り、「差分プライバシー」は特定のプライバシー保護手法を指すものではなく、あくまでその安全性を定義する枠組みである。差分プライバシーに基づきプライバ

[7] なお、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)。

シーを保護する具体的な手法そのもの（定義 1 における $\mathcal{M}$）は「メカニズム（mechanism)」と呼ばれる。差分プライバシーを実現する代表的なメカニズムとしては Laplace メカニズム（後述）が挙げられるが、たとえば PRAM (post randomization) など、一部の伝統的な統計的開示制御 (statistical disclosure control, SDC) の手法（秘匿手法）も差分プライバシーに基づく安全性を与えることが知られている。

3.3　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（問い合わせの種類）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。

たとえば、差分プライバシーを満たす代表的なメカニズムとして知られる Laplace メカニズムを、分割表に適用すること自体は極めて容易である。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）

Laplace メカニズムは、前述の通り 0 を中心とした Laplace ノイズを各セルに加算することにより差分プライバシーを満たす。このノイズは負値もとりうる（半数が負値となる）ことから、人口がゼロないし少人数のセルに対して Laplace メカニズムを適用すると、その

[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。

セルは「負の人口」を取りうる。この非負制約の逸脱は、単にデータとして不自然であるにとどまらず、従来の（非負値の入力を前提とした）分析手法や分析ツールがまったく適用できなくなることを意味することから，実応用の観点でも受け入れ難い。この非負制約の逸脱に対しては、負値をとるセルの値を 0 に補正するという単純な方法により容易に解決可能に思われる。また、この操作により 0 値を持つセルが増加することから、データ量の増大の問題も緩和される。しかし、このような「負値の切り上げ」による単純な解決は、部分和に大きな誤差（過大バイアス）を生じさせる。

また、国勢調査における小地域統計など、詳細な地理的区分を持つ分割表は、しばしば数多くのゼロ値のセルを含む（高いスパース性を持つ)。記憶効率の観点から、疎なデータを計算機上で表現する際にはゼロ値のセルを省略した表現とすることが多いため、そのデータ量は（実際のメッシュ数ではなく）非ゼロ値のセル数にほぼ比例する。しかし、Laplace メカニズムにより加算される Laplace ノイズが 0 となることは確率的にほぼありえないため、Laplace メカニズムの出力はそのすべてが非ゼロ値となり、データ量の著しい増大を招くことになる。

最後に部分和精度の劣化の問題について説明する。たとえば商圏分析などの小地域統計の実応用では、最小単位となるセル値そのものではなく、分析ごとに設定した範囲（たとえばある店舗の商圏など）に含まれる複数セルの合算値（部分和）が重要となる。Laplace メカニズムでは、部分和に加わるノイズは、合算対象となったセルに加わるノイズの総計となり、その分散は、合算の対象となるセルの数に比例する。すなわち、合算対象となる範囲（部分和をとる範囲）が広くなるほど、その部分和の誤差は増大することになる。

なお、国勢調査の統計表など、値として必ず整数値をとる統計表に対しては、(Laplace 分布を離散化した形状を持つ）両側幾何分布に従う乱数（両側幾何ノイズ）を用いる、幾何メカニズム（geometric mechanism）を Laplace メカニズムの代わりに用いることができる。幾何メカニズムは、その出力が必ず整数となるという特徴を持つが、それ以外の性質は Laplace メカニズムとほぼ同様であり、上記の議論もそのまま当てはまる。

3.4　国勢調査に適用可能な差分プライバシーの実現手法

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 Wavelet 変換を用いる方式が有用であることが示されている。Privelet 法 (Xiao et al. (2011)) は、ノイズ付与の過程で Haar Wavelet 変換を導入し、近隣セル間のノイズが相互に打ち消し合う効果をもたらすことにより、連続した領域における部分和の精度を改善する。しかし、その代償として Privelet 法は Laplace メカニズムよりも強いノイズの付与を必要とし、さらに非負制約の逸脱やスパース性の喪失に関する課題を解決しない。Laplace メカニズムと同様に、負値のセルを 0 に切り上げることにより非負制約の逸脱を解決することは可能であるが、やはり同様に過大バイアスの影響を受ける。

寺田他(2015) は、Morton 順序写像と非負精緻化を伴う Wavelet 変換に基づく手法 (本稿では非負ウェーブレット法と呼ぶ) を提案している。この手法では、Privelet 法と同様に Wavelet 空間上でノイズ付与を行なうが、メッシュ人口データなどの二次元データを扱うにあたり、局所性保存写像の一種である Morton 順序写像を用いて一次元データに変換した上で Wavelet 変換を施すことにより、多次元 Wavelet の利用に伴うノイズ強度の増大を防いでいる。また、出力結果が非負制約を逸脱しないように係数値を補正 (非負精緻化) しながら逆 Wavelet 変換を施すことにより、非負制約を充足しつつ差分プライバシーが保証された人口データを得る。非負ウェーブレット法により得られた人口データは、Wavelet 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スパース性も復元されるという特徴を持つ。すなわち、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を用いた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。そのため、非負ウェーブレット法は国勢調査におけるメッシュ統計への差分プライバシーの適用にあたって有力な手法と考えられるが、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。

また、異なるアプローチとして、制約つき最適化を用いる方法が挙げられる。具体的には、Laplace メカニズムや幾何メカニズム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。

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。また、2020 年センサスでは、商用ソルバの Gurobi を用いてこれを実装している。ただし、この制約つき最適化は大きな計算コストを必要とし、たとえば Lee et al. (2015) の手法は 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 年センサスの実装にあたっては、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]。

寺田他 (2017) 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

---

[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。

空間における正規単体への射影問題に帰着されることを利用して、大規模なデータに対して高速に適用可能な手法を示している。この手法は当時の一般的なノート PC を用いて 100,000 セルのデータに対する処理を 12.6 ミリ秒で完了する[12]ことが示されており、国勢調査などの大規模な統計に適した手法と言える。

これらの手法を国勢調査の統計表に適用するにあたっては、一般には最小集計区分（国勢調査の場合は基本単位区）単位の人口のみ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。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し…と、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[13]。

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[14]。

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。特にトップダウン構成法における部分和精度について、理論的には Laplace メカニズムよりも良いことが期待されるが、それがどの程度のものになるか、定量的には明らかになっていない。

3.5　本研究における研究課題と目標

ここまで議論した、差分プライバシーの各実現手法の国勢調査への適用可能性について表 1 にまとめる。Laplace メカニズムや幾何メカニズム、および Privelet 法は、本節で示した課題を解決しないため、統計の有用性に課題がある。非負ウェーブレット法は、これらの課題をすべて解決し、国勢調査データを用いた実証を通じ、特に部分和精度について最も優れることが定量的に示されているが、その一方でメッシュ統計以外へはそのまま適用できず、適用のための修正方法も自明ではない。制約つき最適化を用いた手法はメッシュ

[12] CPU に intel Core i5-6200U (2.3GHz, 2 core)を用いた計測結果（寺田他 (2017)）

[13] 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。

[14] 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。

統計以外への適用も理論的に可能であり、特にトップダウン構成法はすべての課題を解決することが期待されるが、メッシュ統計以外での実証はなされておらず、その実用性や定量的な性質については明らかではない。

表 1: 国勢調査への各手法の適用可能性

	非負制約	スパース性	部分和精度	メッシュ統計	メッシュ統計以外
Laplace メカニズム / 幾何メカニズム	×	×	×	実証ずみ	未実証
Privelet 法	×	×	△	実証ずみ	― (適用不可)
非負ウェーブレット法	○	○	◎	実証ずみ	― (適用不可)
制約つき最適化 (ボトムアップ構成法)	○	○	×	実証ずみ	未実証
制約つき最適化 (トップダウン構成法)	○	○	○	実証ずみ	未実証

そこで、本研究では、2015 年の国勢調査の個票データに基づき、基本単位区を最小の集計区分とする小地域集計を対象とし、制約つき最適化に基づく方法を中心とした差分プライバシーの実現手法の、わが国への国勢調査への適用可能性を実証する。本実証では、以下の事項を研究課題として設定し、これらに関する知見を得ることを目標とした。

- **研究課題 1: 差分プライバシーの適用により、どの程度の誤差がもたらされるか。**
  差分プライバシーの適用によりもたらされる誤差は、使用するメカニズムの種類と安全性（プライバシー損失 ε の値）により変化するため、一概に「差分プライバシーの適用による誤差」がいくつかを論ずることはできない。そこで、安全性をいくつかの段階に変化させた上で、手法ごとにその有用性の変化を定量的に評価する。
- **研究課題 2: 伝統的な手法と比べ、統計の有用性はどの程度変化するか。**
  差分プライバシーは伝統的な手法と比較して、得られるデータの有用性が優れている、もしくは劣っている、との議論がされることがあるが、一般に安全性レベルを下げれば有用性が上がるものであるため、この議論は安全性を揃えて議論しないと意味がない。そこで、 安全性（ε の値）を揃えた上で、伝統的手法の一つである PRAM の適用結果との間で、その有用性について比較評価する。
- **研究課題 3: 差分プライバシーをわが国の国勢調査に適用する際に適した手法は何か。**
  前述の通り、安全性と有用性のバランスはプライバシー保護手法により異なる。また、有用性の観点でどのような統計量が重要かは統計の種類や用途によって異なる。特に、国勢調査における小地域集計のような詳細な地理区分を備えるデータの場合、個々の最小集計単位 (基本単位区など) の値をそのまま利用するだけでなく、それらを複数集約した部分和として用いられることが多い。そこで、部分和をとる範囲のレベルを変えつつ、それぞれの手法の適用結果について有用性を定量的に評価する。

４．　2015 年の国勢調査データへの差分プライバシーの適用

本実験では、前節で示した 3 つの研究課題に関して知見を得ることを目標とし、国勢調査データへの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、およびその実験結果を示す。

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、属性に関する集計区分が異なる（人口のみ、男女別、男女別・年齢 5 歳階級別の 3 種類)。つまり、①基本単位区別人口、②基本単位区別・男女別人口、③基本単位区別・男女別・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、基本単位区別・男女別人口を集計表 2、基本単位区別・男女別・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。

これらの集計表の作成は、いずれもオンサイト施設の環境において、Python による独自プログラムを用い、圧縮効率・アクセス速度に優れるデータ格納形式である Parquet 形式への変換を通じて行った。なお、これらのデータ作成において、オンサイト施設の標準環境ではメモリが不足したため、使用可能メモリを標準の 2 倍 (32GB) に増加させた仮想マシンを利用した。

４.２　実験の方法

本実証では、前述の 3 つの研究課題に対する知見を得ることを目標とし、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、第 3 節での議論から、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 (2017)) 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。つまり、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。

なお、実験で利用した計算機環境（オンサイト施設）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。つまり、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[15]した上で、最後にそれらの結果を統合して評価指標（後述）を計算している。

また、これら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するものとした。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。誤差指標としては絶対平均誤差（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。つまり、$3 \times 4 \times 8 \times 4 = 384$ 種類の実験バリエーションに対してそれぞれ MAE と RMSE を算出したものが、本実験の結果として得られることになる。

4.3　実験の結果

前節で示した 384 種類の実験バリエーションについて、表 2〜4 に MAE を指標とした評価結果を、付表 1〜3 に RMSE を指標とした評価結果を示す。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、表 2、表 3、表 4 はそれぞれ集計表 1、集計表 2、集計表 3 に関する評価結果をまとめたものであり、付表 1〜3 も同様である。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

[15] 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。

合で大きな傾向差が見られなかったため、以降では主に表 2〜4 に基づき MAE を指標として議論する。

なお、表 2 と付表 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出力でも保存する性質を持つ（総数制約の充足）。また、前述の通り本実験では（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（つまり総数）の誤差は 0 となる。つまり、表 2 と付表 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、たとえば日本全国を最上位の地域区分として集計する場合には、この性質を得ることはない（ただし、その代わりに日本全国の人口総数の誤差が 0 となる）。

また、特に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 の適用結果に関する解釈にも留意する必要がある。たとえば表 4 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（$\varepsilon = 0.1$〜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しまっていることを反映している。つまり、表 4 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、ランダムな一様分布に従う出力であったとしても、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 4 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。

表 2 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 3 集計表 2 (基本単位区別・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 4 集計表 3（基本単位区別・男女別・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

5. 実験結果に関する考察

前節では、第 3 節で提示された 3 つの研究課題を明らかにするために、国勢調査の個票データを用いて差分プライバシーの適用可能性に関する実証実験を行った。本節では、前節の実験結果に基づき、上記の研究課題について考察を行う。

第 1 の研究課題は、差分プライバシーの適用によって、どの程度の誤差がもたらされるかを明らかにすることである。

本実験結果は、まず「適用する差分プライバシーの実現手法とプライバシー損失予算（ε）の値によって、もたらされる誤差はまったく異なる」ことを示している。第 3 節で議論した通り、差分プライバシーは、プライバシー損失予算が同一であれば、その実現手法がどのように異なっても同じレベルの安全性を保証するが、その有用性は手法や応用ごとに異なるとされる。本実験で得られた結果は、これを裏付けるものであり、差分プライバシーの適用に関する実用性や有用性を議論するにあたっては、本実験のように適用する手法とプライバシー損失予算の値を変化させつつ、その傾向を定量的に捉えて議論する必要がある。

また、最小の集計区分（本実験では基本単位区ごと）における誤差の傾向と、それらを畳みあげた値（たとえば市区町村ごとなど、より大きな地域区分における部分和）における誤差の傾向が同じとは限らないことも、これらの結果から確認できる。具体的には、基本単位区レベルでの誤差のみを有用性の指標とすると、ボトムアップ構成法や Laplace メカニズムを適用した結果が優れているように見える[16]。しかしながら、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人口においては、ボトムアップ構成法と Laplace メカニズムのいずれについても誤差が拡大する傾向にあり、特に Laplace メカニズムではそれが顕著である。これは、本実験における Laplace メカニズムの適用における、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ためである[17]。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。

それに対して、トップダウン構成法については、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。具体的には、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

[16] 表 3 において ε が小さい場合は PRAM が最も優れているように見えるが、これは前節で議論した通り「見せかけ」の精度であり、実質的な意味を持たない。

[17] つまり、負値の切り上げを行わなければここまで誤差が拡大することはないと言えるが、そのかわりに非負制約も充足されなくなるため、結果表に数多くの「負の人口」を含みうることになる。

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。ただし、使用する変数が増大するにつれて、また分類区分が詳細になるにつれて、より多くの数のセルを備えた集計表が作成されるが、そのような集計表に含まれるセルを合算した場合、部分和の誤差はより増大する傾向にあることから、結果数値を比較する上では、注意が必要である。

第 2 の研究課題は、伝統的な匿名化手法である PRAM に比べ、結果数値の有用性がどの程度異なるかを確認することである。

前述の通り、差分プライバシーにおいては、プライバシー損失予算の値が同じであれば、その実現手法がいかに異なるものであっても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。換言すると、手法ごとの「プライバシー保護の効率」の違いを比較できる。同一の安全性条件下で誤差が小さいほど（すなわち有用性の低下の度合いが小さいほど)、同じ安全性を達成する上での有用性に関する犠牲が小さい、つまりプライバシー保護の効率に優れていることを意味する。

この観点で PRAM を他の手法と比較すると、同一条件下では 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。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、これは第 4 節で議論した通り「見せかけ」の精度であり、なんら実用的な意味を持たない。

さらに、PRAM 以外の手法は、ε の値を大きくするにつれ有用性が向上するが、PRAM ではその傾向がほとんど見られない。これは、PRAM のプライバシー保護の効率が他の手法に比べて「悪すぎる」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、ρ の値はほぼ 0 に近い値になることが多い。つまり、ε の値を多少大きくして安全性を緩和しても ρ は 0 に近い値に留まったままであり、有用性の改善には繋がらない。このことは、伝統的な秘匿処理手法の一つである PRAM は、差分プライバシーを満たすことができるが、そのプライバシー保護の効率は極めて悪く、他の手法を検討することが望ましいことを示唆している。

第 3 の研究課題は、差分プライバシーをわが国の国勢調査に適用する際に適した手法は何かということである。

研究課題 1 に関する議論で示した通り、単純な Laplace メカニズムは非負制約の解決が

課題となり、たとえば本実験で実施したように単純に負値の切り上げにより非負制約を解決しようとしても、部分和における大きな過大推計バイアスの影響により実用に耐える集計表を得ることは困難である。また、研究課題 2 の議論から、PRAM はそのプライバシー保護効率の悪さから、適切な安全性と有用性を両立させることは難しいことが示唆される。

ボトムアップ構成法とトップダウン構成法を比較・検討すると、基本単位区レベルの誤差を見ると、ボトムアップ構成法の有用性が相対的に高いが、部分和で見た場合の誤差に関しては、部分和の範囲が大きくなるにつれて、ボトムアップ構成法の誤差は相対的に増大し、有用性が小さくなる。それに対して、トップダウン構成法においては部分和に伴う誤差の程度が抑えられるため、上位の地理的区分に合わせて部分和を算出した場合でも、有用性が保持されることがわかった。

このことから、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有用性の高い技法であるが、部分和については誤差が大きく増大する傾向にあるため、有用性が小さくなることが確認できた。トップダウン構成法については、基本単位区レベルにおける誤差はボトムアップ構成法のそれよりも大きくなるが、安全性の水準が適切に設定されれば、今回の対象となる変数の組み合わせにおける結果数値の有用性は、部分和を考慮しても保持されていることが確認できた。

ただし、本実証実験で適用された変数群は限定されたものであるだけでなく、有用性の評価指標としては、部分和を含む対象となるセルに関する結果数値の誤差のみが用いられており、それ以外の統計量は評価指標に含まれていない。したがって、国勢調査への適用可能性を追究する上では、各種の統計量を用いたさらなる検証を行うことが求められよう。

6. おわりに

本稿では、差分プライバシーの方法論の国勢調査へ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作成した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手法を適用した場合の有用性を評価するだけでなく、伝統的な匿名化手法の 1 つである PRAM と比較・検証を行った。本研究の結果から、わが国の国勢調査に差分プライバシーを適用する場合、トップダウン構成法における有用性が、他の手法と比較して相対的に高いことが確認できた。本稿で述べたように、センサス局では、2020 年センサスにおける詳細統計表の作成・公表にあたって、トップダウンアルゴリズムの一種である TDA を採用していることから、本研究成果は、センサス局の実務で適用された技法にも符合していると言えよう。このことは、地理的区分の階層的な構造においては、基本単位区における集計表のセルにノイズを付与した上で、小地域区分から畳み上げて、上位の地理的区分において集計を行うよりも、差分プライバシーに基づいて発生するノイズ全体をトップダウンで整合を取りつつ統計表の各セルに割り当てる方が、有用性の観点から見れば合理

的な結果をもたらすことを意味している。

その一方で、本研究は、差分プライバシーの方法論をわが国の国勢調査における基本単位区別の詳細な集計表に適用した最初の事例であることから、試論的な側面を持っている。本研究では、基本単位区別で見た年齢と性別のクロス表を用いた検証を行ったが、国勢調査では、年齢や性別以外の人口社会的な調査項目、就業状況、産業や職業といった就業特性に関する調査項目、さらには住宅に関する調査項目が捕捉されており、これらの項目を集計事項として作成した統計表に対して差分プライバシーの実現手法を用いた場合の有用性の検証は、今後の検討の対象となりうる。こうした様々な調査項目に基づいて作成される各種の集計表を対象に、差分プライバシーの有効性をさらに追究することについては、今後の検討課題としたい。

参考文献

[1] 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.

[2] 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.

[3] 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.

[4] 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 号において刊行予定)

[5] 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.

[6] 寺田雅之・山口高康・本郷節之(2017)「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.

[7] 寺田雅之 (2018)「差分プライバシーとは何か」『システム/制御/情報』 第 63 巻第 2 号, システム制御情報学会, pp.58-63.

[8] 寺田雅之 (2019)「差分プライバシの基礎と動向」『情報処理』 第 61 巻第 6 号, 情報処理学会, pp.591-599.

[9] Abowd, J. M. (2021) “Supplemental Declaration of John M. Abowd.”, Alabama v. U.S. Dep’t of Commerce, 546 F. Supp. 3d 1057 (M.D. Ala. 2021).

[10] Bach, F. (2022) “Differential Privacy and Noisy Confidentiality Concepts for European

Population Statistics", *Journal of Survey Statistics and Methodology* (10), pp.642-687.

[11] Burwell, S. M., VanRoekel, S., Park, T., and Mancini, D. J. (2013) *Open Data Policy – Managing Information as an Asset*, Memorandum M-13-13. Office of Management and Budget, Executive Office of the President, U.S.

[12] Dove, I. (2021) "Applying differential privacy protection to ONS mortality data, pilot study".

[13] Duncan, G. T., Elliot, M., Salazar-González, J-J.(2011) Statistical Confidentiality, Springer. Dwork, C. (2006) "Differential privacy", Proc. 33rd intl. conf. Automata, Languages and Programming, LNCS 4052, Springer, pp. 1-12.

[14] 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.

[15] Fraser, B. and Wooton, J. (2005). A proposed method for confidentialising tabular output to protect against differencing, Paper Presented at Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality,Geneva, Switzerland.

[16] 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.

[17] Garfinkel, S. (2022) "Differential Privacy and the 2020 US Census", MIT Case Studies in Social and Ethical Responsibilities of Computing, Winter 2022.

[18] Heldal, J., Johansen, S., Risnes, Ø. (2019) "Instant Access to Microdata – microdata.no", Paper presented at New Techniques and Technologies for Statistics 2019, Brussels.

[19] 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.

[20] 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.

[21] McKenna, L. (2019) "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] Office for National Statistics (2017) "Development of flexible dissemination for 2021 Census".

[23] Smith, H. Jeff, Sandra J. Milberg, and Sandra J. Burke (1996) "Information Privacy: Measuring Individuals' Concerns About Organizational Practices." *MIS Quarterly* 20(2): 167-196.

[24] Xiao, X., Wang, G., Gehrke, J. and Jefferson, T. (2011). Differential Privacy via Wavelet Transforms, *IEEE Transactions on Knowledge and Data Engineering*, **23**(8), 1200–1214.

付録

付表 1 集計表 1（基本単位区別人口）に対する評価結果（RMSE）

ε	手法	RMSE(都道府県)	RMSE(市区町村)	RMSE(町字)	RMSE(基本単位区)
0.1	(a)PRAM	0.00	27444.13	1199.16	91.17
	(b)Laplace	132959.85	4810.29	152.35	25.21
	(c)BottomUp	0.00	1497.94	120.57	24.83
	(d)TopDown	0.00	110.71	104.19	73.91
0.2	(a)PRAM	0.00	27466.61	1199.82	91.18
	(b)Laplace	41949.58	1555.75	67.33	13.24
	(c)BottomUp	0.00	620.28	60.68	13.19
	(d)TopDown	0.00	58.57	53.24	43.32
0.7	(a)PRAM	0.00	27444.77	1198.52	91.18
	(b)Laplace	7594.78	294.57	18.38	3.96
	(c)BottomUp	0.00	153.61	17.75	3.96
	(d)TopDown	0.00	16.24	15.92	14.75
1	(a)PRAM	0.00	27462.95	1198.92	91.18
	(b)Laplace	4834.71	195.62	12.67	2.76
	(c)BottomUp	0.00	107.37	12.28	2.76
	(d)TopDown	0.00	10.98	11.10	10.46
1.1	(a)PRAM	0.00	27446.66	1199.26	91.18
	(b)Laplace	4358.03	172.91	11.56	2.51
	(c)BottomUp	0.00	93.66	11.19	2.51
	(d)TopDown	0.00	10.12	10.09	9.55
5	(a)PRAM	0.00	27376.10	1195.51	90.93
	(b)Laplace	825.62	35.24	2.54	0.55
	(c)BottomUp	0.00	20.39	2.46	0.55
	(d)TopDown	0.00	2.24	2.22	2.17
10	(a)PRAM	0.00	21077.24	909.39	69.94
	(b)Laplace	444.96	17.79	1.26	0.28
	(c)BottomUp	0.00	10.15	1.24	0.28
	(d)TopDown	0.00	1.16	1.11	1.09
20	(a)PRAM	0.00	7.38	0.67	0.14
	(b)Laplace	207.11	8.70	0.63	0.14
	(c)BottomUp	0.00	5.02	0.62	0.14
	(d)TopDown	0.00	0.58	0.56	0.54

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

ε	手法	RMSE(都道府県)	RMSE(市区町村)	RMSE(町字)	RMSE(基本単位区)
0.1	(a)PRAM	50695.95	13893.39	603.19	46.34
	(b)Laplace	211871.50	7437.22	190.38	23.43
	(c)BottomUp	6177.66	1727.16	115.55	22.41
	(d)TopDown	78.90	112.18	99.33	59.15
0.2	(a)PRAM	49727.26	13889.06	603.07	46.34
	(b)Laplace	68052.87	2435.92	76.67	12.59
	(c)BottomUp	3105.15	775.52	60.33	12.41
	(d)TopDown	32.35	55.43	51.88	36.88
0.7	(a)PRAM	43181.26	13857.27	602.94	46.33
	(b)Laplace	9703.96	363.85	18.90	3.91
	(c)BottomUp	640.22	166.42	17.81	3.90
	(d)TopDown	11.76	16.58	15.80	13.68
1	(a)PRAM	39442.99	13847.90	602.63	46.33
	(b)Laplace	5924.25	226.16	12.88	2.73
	(c)BottomUp	408.21	111.27	12.28	2.73
	(d)TopDown	10.10	11.32	11.02	9.91
1.1	(a)PRAM	38049.07	13834.08	602.71	46.33
	(b)Laplace	5102.39	198.19	11.68	2.49
	(c)BottomUp	370.07	100.62	11.17	2.49
	(d)TopDown	7.72	10.22	10.03	9.10
5	(a)PRAM	7945.16	13761.60	602.20	46.32
	(b)Laplace	908.66	35.63	2.55	0.55
	(c)BottomUp	85.58	20.14	2.47	0.55
	(d)TopDown	1.58	2.19	2.22	2.14
10	(a)PRAM	736.38	13736.09	600.53	46.20
	(b)Laplace	443.94	18.19	1.28	0.28
	(c)BottomUp	35.84	10.56	1.24	0.28
	(d)TopDown	0.70	1.16	1.11	1.08
20	(a)PRAM	10.18	10563.20	456.81	35.56
	(b)Laplace	222.95	8.83	0.63	0.14
	(c)BottomUp	18.92	5.08	0.62	0.14
	(d)TopDown	0.33	0.56	0.56	0.54

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

ε	手法	RMSE(都道府県)	RMSE(市区町村)	RMSE(町字)	RMSE(基本単位区)
0.1	(a)PRAM	28567.92	1197.12	43.51	3.79
	(b)Laplace	499151.33	16450.12	357.83	20.10
	(c)BottomUp	24372.95	1058.46	50.55	8.67
	(d)TopDown	116.38	107.55	56.97	13.55
0.2	(a)PRAM	28529.57	1195.88	43.50	3.79
	(b)Laplace	233353.87	7732.67	169.09	10.15
	(c)BottomUp	19974.50	865.28	37.55	5.94
	(d)TopDown	58.30	54.62	34.56	10.11
0.7	(a)PRAM	28235.19	1189.69	43.42	3.79
	(b)Laplace	53698.26	1806.96	40.64	3.09
	(c)BottomUp	8905.23	387.35	16.29	2.58
	(d)TopDown	16.09	15.99	12.59	5.66
1	(a)PRAM	27990.29	1184.62	43.33	3.79
	(b)Laplace	33930.09	1148.94	26.28	2.20
	(c)BottomUp	6429.64	279.62	11.93	1.94
	(d)TopDown	10.92	11.11	9.17	4.63
1.1	(a)PRAM	27920.58	1182.51	43.31	3.79
	(b)Laplace	30132.23	1021.71	23.48	2.02
	(c)BottomUp	5867.40	254.39	10.96	1.80
	(d)TopDown	10.18	10.23	8.44	4.39
5	(a)PRAM	23131.76	1088.11	42.07	3.78
	(b)Laplace	4781.46	165.45	4.17	0.49
	(c)BottomUp	1236.76	54.73	2.64	0.48
	(d)TopDown	2.28	2.24	2.08	1.54
10	(a)PRAM	11450.98	916.52	40.00	3.76
	(b)Laplace	2301.65	79.79	2.04	0.25
	(c)BottomUp	615.17	27.18	1.34	0.25
	(d)TopDown	1.06	1.12	1.07	0.87
20	(a)PRAM	644.74	855.66	38.97	3.71
	(b)Laplace	1148.93	39.83	1.02	0.13
	(c)BottomUp	307.43	13.64	0.67	0.12
	(d)TopDown	0.55	0.56	0.54	0.46