

国勢調査を対象にした差分プライバシーの 有用性評価に関する一考察 —生態学的回帰に着目して—

2026年9月9日

(独)統計センター/総務省統計研究研修所/中央大学 加藤 駿典

中央大学 伊藤 伸介

NTTドコモ/京都橘大学 寺田 雅之

滋賀大学/京都橘大学 松井 秀俊

本報告の内容は個人的見解を示すものであり、
統計センターの見解を示すものではありません。

- 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では、**攪乱的手法(perturbative methods)**が積極的に採用されてきた(伊藤・寺田(2023))。
- **データベース再構築攻撃(Database Reconstruction Attack)**が現実的な脅威として指摘された(Abowd(2018))。
- その対策として、アメリカセンサス局では、**差分プライバシー(Differential Privacy)**の適用を検討し始めた(Jamin(2021))。

アメリカセンサス局は、

- 2010年人口センサスを用いて差分プライバシーの実用性に関する検証を行った。
- その際、実現方式として**TopDownアルゴリズム(TDA)**を用いた。
 - 統計表の公表で消費されるプライバシー損失予算 ϵ を地域レベルにどう配分するかについての調整も行った (Garfinkel et al.(2019))。

↓

- 2020年人口センサスを対象に、差分プライバシーを適用した統計表
区画改定データ(PL94-171) やPPMFs(=Privacy-Protected
Microdata Files)を公開した(伊藤他(2022), 伊藤他(2024))。

プライバシー保護における「安全性指標」を与える枠組み

- 差分プライバシーの実現方式は数多く存在。
(集計値やデータにノイズを付加する方法が一般的)
- 様々な方式の安全性を、プライバシー損失予算 (privacy loss budget) “ ϵ ” という統一的な指標で測ることができる(値が小さいほど安全)。

任意の攻撃能力や背景知識を持つ攻撃者に対する安全性

「想定外」の攻撃に対する堅牢性 (どんな AI ができても安全)

→ 任意のデータの組み合わせに対する再構築攻撃にも安全。

【定義】

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 $\mathcal{A}: \mathcal{D} \rightarrow \mathcal{R}$ が下記を満たすとき、 $\mathcal{A}$ は ε -差分プライバシーを満たす。

ただし、ここで S は $\mathcal{A}$ の出力空間 $\mathcal{R}$ の任意の部分集合である ($S \subseteq \mathcal{R}$)。

$$\Pr[\mathcal{A}(D_1) \in S] \leq e^\varepsilon \times \Pr[\mathcal{A}(D_2) \in S]$$

Aさんの情報が入っているDBへの問い合わせ結果と、
Aさんの情報が入っていないDBへの問い合わせ結果が区別できないなら、
この問い合わせにより **Aさんのプライバシーが漏れることはない。**
⇒差分プライバシーを満たす。



$$\Pr[\mathcal{A}(D_1) \in S] \leq e^\epsilon \times \Pr[\mathcal{A}(D_2) \in S]$$

D1 への問い合わせ結果の確率分布

D2 への問い合わせ結果の確率分布

- わが国でも、国勢調査データを対象とした差分プライバシー(DP)の適用可能性に関する実証研究が進められてきた(伊藤他(2024), 伊藤他(2025), 伊藤他(2026))。
- 伊藤他(2025)は、生態学的誤謬(ecological fallacy)の観点から(Robinson(1950))、DPを適用して作成した**フルクロス表**と**按分クロス表**の有用性を比較検証した。
 - **フルクロス表**:対象となるすべての変数を用いて作成したクロス表
 - **按分クロス表**:上位の地域区分から按分を適用して作成した擬似的なクロス表

⇒これらの研究では、DPが適用されたデータの有用性評価に、セル単位の誤差を測る平均絶対誤差(MAE)や二乗平均平方根誤差(RMSE)が主に用いられてきた。

- 伊藤他(2024)は、MAEやRMSEでは、DPによるノイズ付加の影響を十分に評価できない場合があることを指摘している。
→代わる評価方法として、**Cohen et al.(2022)**の研究に注目している。
- Cohen et al.(2022)は、「**生態学的回帰(ecological regression)**」※に着目し、TDAの有効性を検証している。
※地域レベルの集計データに対する回帰分析により、個人レベルの傾向を推定する方法(Goodman(1953))
→2010年アメリカ人口センサスに基づき、人種・民族と投票行動に関する生態学的回帰を行い、DP適用前後で回帰係数がどれほど変化するか検証した。その結果、重み付き回帰(**WLS**)では、大きな変化が見られないことが確認された。

- 本報告では、Cohen et al.(2022)の評価方法をわが国の国勢調査データに適用することで、生態学的回帰による有用性評価に関して検討する。
- 具体的には、令和2年国勢調査の個票データから作成した集計表にDPを適用し、DP適用前後で生態学的回帰の係数がどの程度変化するかを評価したうえで、この評価方法の有効性について議論する。
- 本分析では、被説明変数を持ち家率、説明変数を人口密度として、労働力状態による層別回帰とする。

- 令和2年国勢調査の個票データ (調査票情報) に基づいて作成したクロス表
- **使用する変数:**
 - **持ち家の有無:**「住居の種類・住宅の所有の関係(8区分)」のうち “持ち家” に該当する場合を1、それ以外を0と設定。
 - **世帯主の続き柄:**「世帯主との続柄(12区分)」のうち “世帯主または代表者” に該当する場合を1、それ以外を0と設定。
→世帯単位での持ち家率の算出が可能。
 - **労働力状態:**就業者、完全失業者、非労働力人口の3区分
※使用する変数のいずれかが不詳か、対象外(施設等の世帯員)に該当するレコードは除外。
- **人口密度の算出に用いる面積データ**は、令和2年国勢調査の境界データを使用。
- **地域区分:**全国、都道府県、市区町村、町字と基本単位区の5区分

- クロス表に対して、Laplaceメカニズムに基づくトップダウン構成法によりDPを適用する。
- プライバシー損失予算 ϵ : **0.1, 0.2, 0.7, 1**
 - ϵ の配分については、地域区分ごとに均等に行う。
(隣接条件については、秘匿n条件 (Unbounded DP) とする。)
- 集計方法:
 - フルクロス法: 出力対象の地域区分※におけるすべての属性を含むクロス表を直接作成する方法
 - 按分法:
 - (1) 1つ上位の地域区分におけるすべての属性を含むクロス表を作成
 - (2) 出力対象の地域区分における属性なしの総計値を作成
 - (3) (1)に基づき(2)を按分することで、出力対象の地域区分のクロス表を推計

※出力対象の地域区分は基本単位区とする。また、町字や市区町村レベルの表については、DP適用済みの基本単位区レベルの表の集約により作成する。

- 国勢調査へのDP適用に際しては、ノイズの付加に加えて、それによって生じる統計表上の矛盾(例:負の人口値など)を解消する後処理も含めた、**実現方式**の選択が重要となる。
- 伊藤他(2024)では、いくつかの**実現方式**を比較検証し、その中で**トップダウン構成法**が最も有効であることを実証的に示した。

トップダウン構成法:

- ノイズ付加と後処理を全国表から「トップダウン」の方向(→都道府県表→市区町村表→町・字表→基本単位区表)に、階層的に適用。
- 2020年アメリカセンサスのTDAも同様のアプローチ (もう少し複雑)

- ①**DP適用前**、DP適用後(②**フルクロス法**、③**按分法**)の集計表それぞれに生態学的回帰を実行し、回帰係数の変化により②③の有用性を評価。
 - 被説明変数:世帯単位の持ち家率
 - 説明変数:人口密度(人/m²) = 全人口/ 面積(m²) ⇒ Box-Cox変換を適用※¹
 - 層別:労働力状態(3区分)による(層別は持ち家率のみ。人口密度は全人口から算出)。
- 推定方法**: Cohen et al. (2022) と同様に、最小二乗法(**OLS**)と加重最小二乗法(**WLS**)※²。
- 対象地域**: 全国
- 地域区分**: 基本単位区、市区町村※³

※¹ λ はDP適用前のデータから最尤推定。DP適用後も前後を同一スケールで比較するために、同一値を用いる(基本単位区:0.222、町字:0.134、市区町村:0.016)。

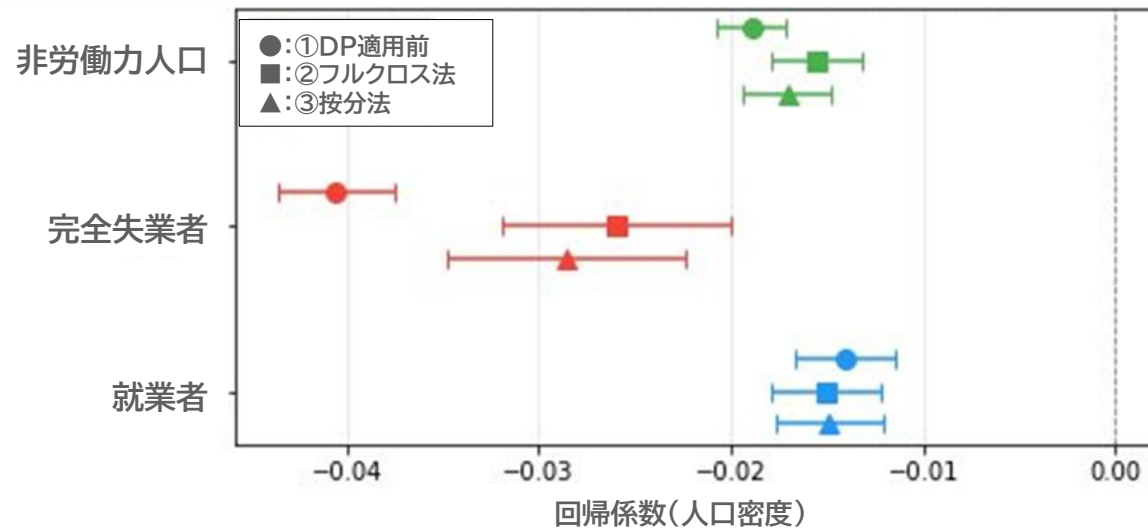
※² ウェイトは、被説明変数の分母である世帯数

※³ 対象地域や地域区分により変数間の関係性が異なると予想されるため、対象地域(全国および4都県)、地域区分(基本単位区、町字、市区町村)ごとに実験を行った。本報告では、そのうち全国の基本単位区・市区町村レベルの結果のみを示す。

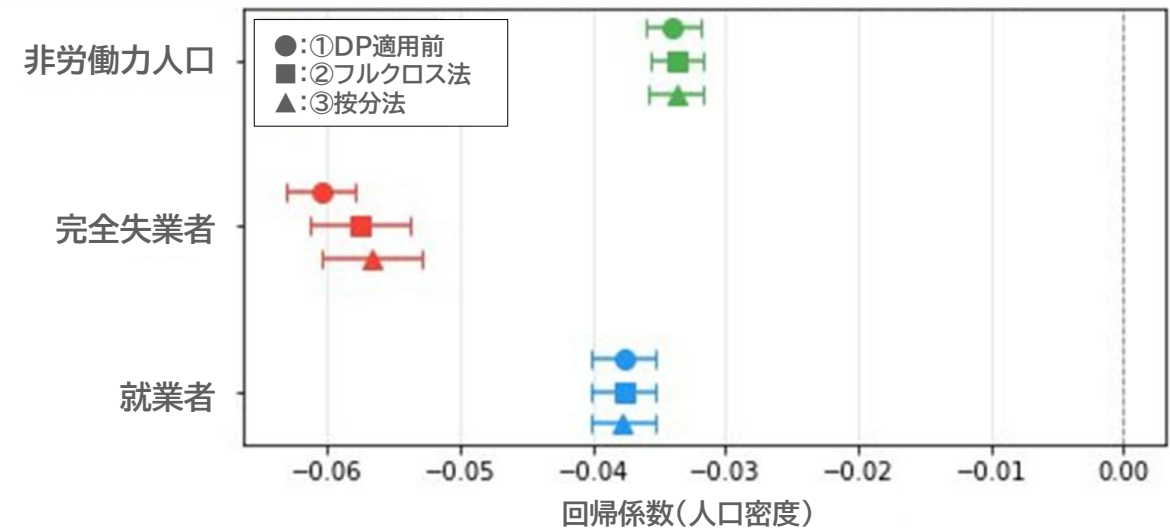
図1 DP適用前後におけるOLSとWLSによる回帰係数の比較($\epsilon=0.1$ 、市区町村別)

- OLSでは①DP適用前と比較して、②フルクロス法と③按分法の回帰係数が大きく乖離する。
- 一方、**WLS**では①と②③の回帰係数の乖離が小さい。

係数プロット(OLS)



係数プロット(WLS)



※エラーバーは回帰係数の95%信頼区間

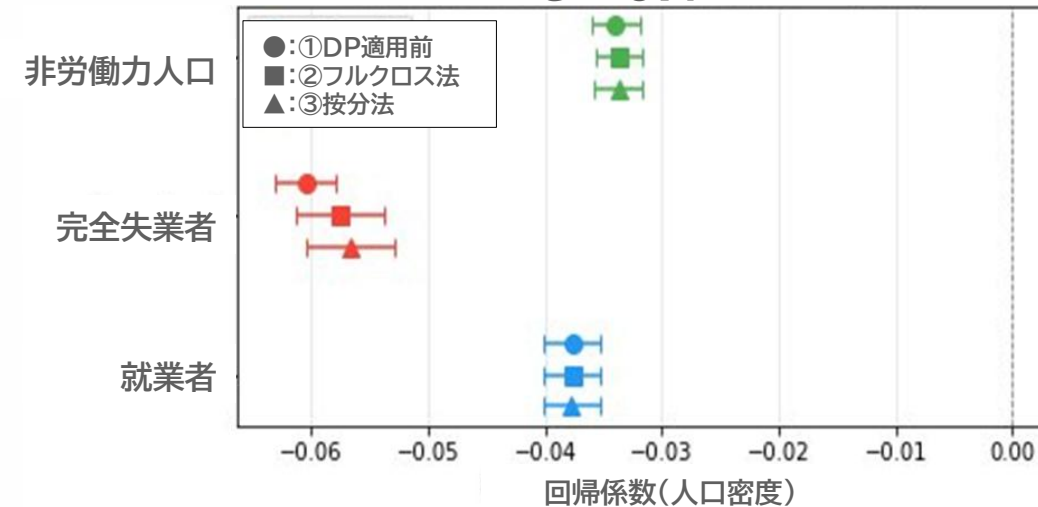
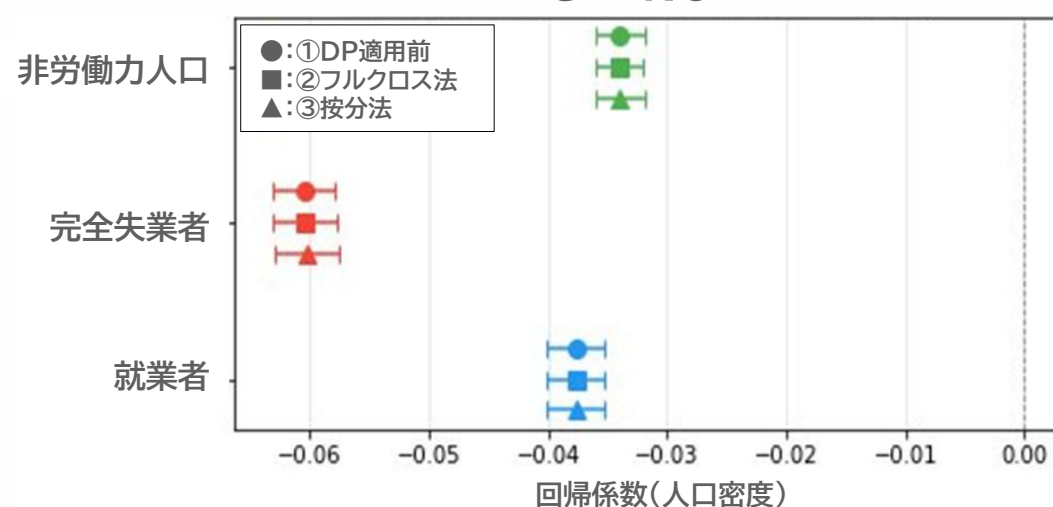
図2 DP適用前後における市区町村別と基本単位区別の回帰係数の比較(WLS)

- ε を0.1まで小さくすると①と②③の回帰係数に差異が見られるが、統計的有意性は保たれている。
- ただし基本単位区レベルでは、完全失業者において①と②③の乖離が大きい。

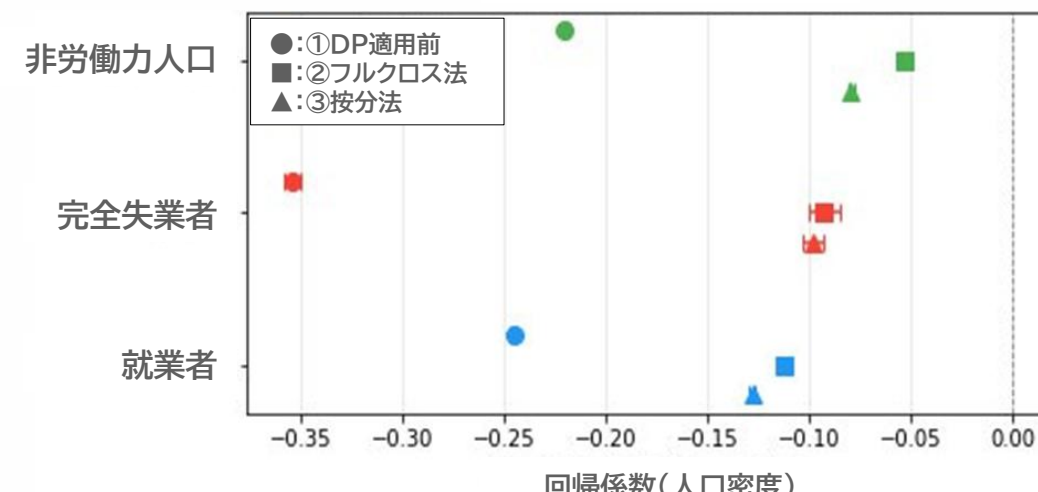
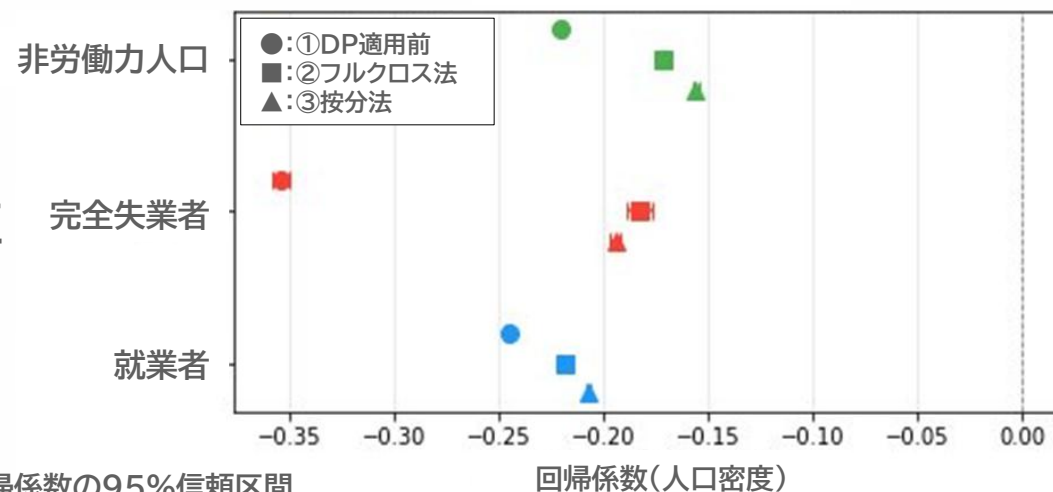
$\varepsilon = 1.0$

$\varepsilon = 0.1$

市区町村



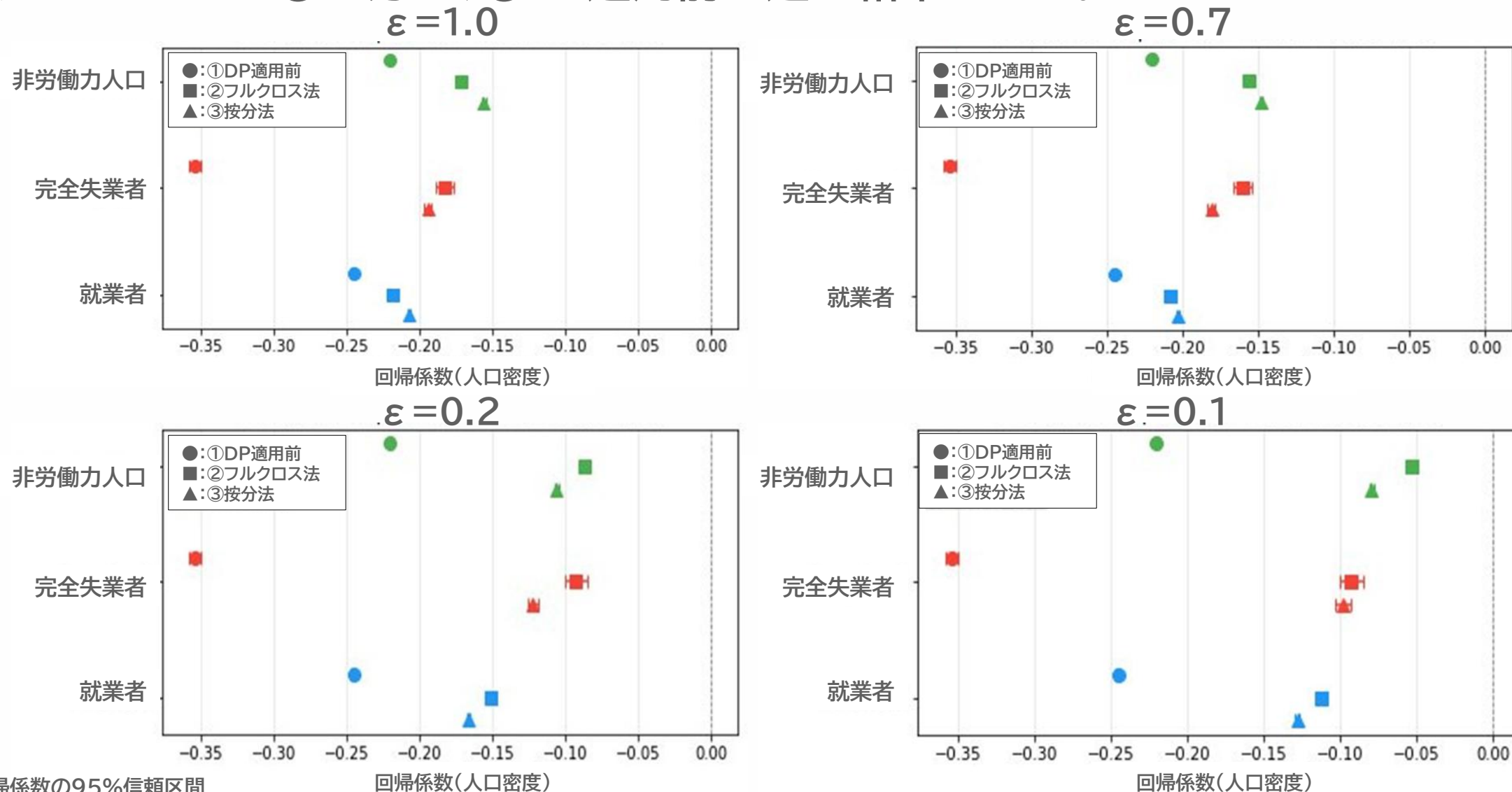
基本単位区



※エラーバーは回帰係数の95%信頼区間

図3 DP適用前後におけるフルクロス法と按分法による回帰係数の比較(WLS、基本単位区別)

- ②フルクロス法と③按分法を比較すると、 ε が大きい場合は②の方が、 ε が小さくなるにつれて③の方が、①DP適用前に近い結果となる。



※エラーバーは回帰係数の95%信頼区間

- 図1より、OLSと比較して、**WLS**の回帰係数はDP適用による影響が小さく、**安定した結果**が得られる。
→Cohen et al.(2022)の重み付き回帰における回帰係数の安定的な傾向と整合的
- 図2より、 $\varepsilon=0.1$ まで小さくしても統計的有意性は保たれており、**DP適用による影響は限定的**である。(Cohen et al.(2022)と同様の結果)
※ただし、基本単位区レベルの完全失業者では影響が大きく推定が不安定となる。これは、当該世帯主が世帯主の約2%と少ないことによると考えられる。
- 図3より、 ε が大きい場合は②フルクロス法が、 ε が小さい場合は③按分法が優位となる。
→伊藤他(2025)の**MAE**による有用性比較の結果と整合的

- 本実験結果はプライバシー損失予算 (ϵ) の配分を地域区分ごとに均等にした結果、得られたものであり、2020年アメリカセンサスのように、この配分の比率を変更した場合には、異なる実験結果が得られると考えられる。
- この配分の比率がもたらす影響が何に起因するかを精密に分析した上で、その影響を最小化するようなメカニズムやpost-processing手法の開発とその適用が有用であると考えられる。

- 本報告では、差分プライバシーの国勢調査への適用可能性にあたっての有用性の評価方法を探究するために、令和2年国勢調査の個票データを用いて作成した集計表をもとに、**DP適用前後の生態学的回帰の係数比較に関する実証実験**を行い、生態学的回帰による有用性評価の有効性を議論した。
- 本実証実験の結果から、我が国の国勢調査データを用いた場合でも、**Cohen et al.の実験と同様に、DP適用後でも生態学的回帰の安定性が確認された。**
- 本報告では、フルクロス法と按分法の比較を行った。 ϵ が大きい場合はフルクロス法が優位、 ϵ が小さい場合は按分法がそれぞれ優位となる傾向が生態学的回帰によって明らかになった。この結果は伊藤他(2025)のMAEによる比較と整合する傾向であった。

- 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, 131～150頁
- 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, 117～136頁
- 伊藤伸介・寺田雅之・加藤駿典 (2024)「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—」『統計研究彙報』第81号, 69～88頁
- 伊藤伸介・寺田雅之・加藤駿典・松井秀俊 (2025)「差分プライベートな国勢調査データの有用性に関する定量的な評価研究」『統計研究彙報』第82号, 101～120頁
- 伊藤伸介・寺田雅之・加藤駿典・松井秀俊(2026)「国勢調査を対象にした差分プライバシーの実用性に関する検討—Zero-Concentrated Differential Privacy (zCDP)に着目して—」『統計研究彙報』第83号, 167～184頁
- Abowd, J. M. (2018). Staring-down the database reconstruction theorem, Joint Statistical Meetings, Vancouver, BC, Canada.
- 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.
- Dwork, C. (2006) "Differential privacy", Proc. 33rd intl. conf. Automata, Languages and Programming, LNCS 4052, Springer, pp. 1-12.
- 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.
- Goodman, L. A. (1953) "Ecological Regression and Behavior of Individuals", American Sociological Review, vol.18, pp.663-664.
- Jamin, R. (2021), "Disclosure Avoidance for the 2020 Census: An Introduction", U.S. Census Bureau.
- Robinson, W. S. (1950) "Ecological Correlations and the Behavior of Individuals", American Sociological Review, vol.15, pp.351-357.

謝辞

本研究は、JST経済安全保障重要技術育成プログラム【JPMJKP24U5】の支援を受けたものです。

- あるデータベースから生成された (一見して安全に見える) データを
重ね合わせることによって制約充足問題を構築し、
- その問題を解いて元のデータベースを復元することにより、データに
含まれる**個人のプライバシーを暴露**する攻撃。

