

## 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

中央大・経済 伊藤 伸介 (株)NTT ドコモ 寺田 雅之

レコードに含まれる属性値に対して、ノイズの付与やスワッピングのような秘匿処理の方法を導入することは、公的統計の匿名化マイクロデータの作成のために用いられるだけでなく、集計結果表の元になる個票データに対しても適用されてきた。例えば、アメリカセンサス局で実施された 2000 年人口センサスにおいて、集計結果表の元になる個票データを作成するために、レコードスワッピングが適用されていたことはその 1 つの例と言える(Zayatz(2007))。アメリカセンサス局が 2010 年のセンサスにおいて統計表や PUMS(=Public Use Microdata Sample)の作成のために用いてきた方法は、情報の削除やトップ(ボトム)コーディング等の非攪乱的手法、およびノイズ付与やスワッピング、さらには部分的な合成データの手法の適用を含む攪乱的手法である。これらの手法を適切に組み合わせることによって、公表可能な統計表や PUMS として提供可能なレベルにまで特定化のリスクが低減される(Lauger et al. (2014))。しかしながら、個体が特定化されないように公表統計表のセルの度数にセル削除等の秘匿処理を施したとしても、同じ統計調査に関して公表された他の統計表群を用いることによる個人情報への漏洩を回避する可能性を保証しない。

アメリカセンサス局は詳細な地域区分を有する統計表を数多く作成しているが、これらの統計表の組み合わせによっては、個票データを公開していなくても、個体が特定されるリスクが高まることが考えられる(Abowd(2018))。これは、database reconstruction attack と呼ばれており、少数のランダムなクエリを組み合わせることによって、クエリの元にある個票データに含まれる個人情報の暴露が可能になる(Dinur and Nissim(2003))。こうしたリスクの回避には、結果数値の精度を考慮しながら、ノイズを付与することによって、安全な統計表を公表することが求められる。そこで、アメリカセンサス局では、Abowd を中心として、差分プライバシーの方法論を統計作成の実務レベルで全面的に導入する一大プロジェクトが進められてきた。

アメリカセンサス局は、2020 年の人口センサスに向けて、2010 年の人口センサスを用いた差分プライバシーの適用可能性に関する実証研究を行っている。全国レベルの性別、人種、年齢、世帯主との続き柄に関する様々な集計結果表を対象に、精度を確保した上で公表することを可能にしつつ、安全性を確保するために、差分プライバシーの実用性に関する検証を行っている。具体的には、州レベル、郡レベル、センサストラックレベル、ブロックレベルで統計表が作成されていることから、地域レベルのそれぞれの段階における統計表の作成のためにプライバシー損失予算(privacy-loss budget)を設定し、プライバシーの損失と統計データの精度のバランスを考慮した上で、プライバシー損失予算に関する最適な値が決定されている。

本報告では、アメリカセンサス局における差分プライバシーの実証実験を参考にしながら、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議論していきたい。

### 参考文献

- Abowd, J. M. (2018) “Staring-Down the Database Reconstruction Theorem”, presented at Joint Statistical Meetings, Vancouver, BC, Canada.
- Dinur, I., Nissim, K. (2003) “Revealing Information while Preserving Privacy” in Proceedings of Twenty-Second ACM SIGMOD-SIGACT-SIGART symposium on Principles of Database Systems (PODS '03), ACM, New York, USA.
- Lauger, A., Wisniewski, B. and McKenna, L. (2014) Disclosure Avoidance Techniques at the U.S. Census Bureau: Current Practices and Research, Research Report Series (Disclosure Avoidance #2014-02), U.S. Census Bureau, pp.1-13.
- Zayatz, L. (2007) “Disclosure Avoidance Practices and Research at the U.S. Census Bureau: An Update”, *Journal of Official Statistics*, Vol.23, No. 2, pp.253-265.