

匿名データの利用改善に向けた調査研究の概要

統計委員会

匿名データ部会に対し検討要請

- ・地域情報などの提供
- ・利用し易いファイル作成 等



匿名データ部会

匿名データ部会（平成28年6月7日開催）

- ・検討課題を確認
- ・検討方法（調査研究で検討）を確定
- ・データ検証作業等は調査研究で

匿名データの作成状況

調査名	リサンプリング率	地域情報
国勢調査	1%	都道府県及び人口50万以上の市区
全国消費実態調査	80%	2区分（3大都市圏か否か）
社会生活基本調査	80%	2区分（3大都市圏か否か）
就業構造基本調査	80%	2区分（3大都市圏か否か）
住宅・土地統計調査	10%	都道府県
労働力調査	80% (沖縄県は20%)	なし
国民生活基礎調査	20%	なし



調査研究

- 国勢調査で実証実験
 - ・都道府県別×性別×年齢（各歳、90歳でトップコーディング）においても最小度数は500以上
 - ・5万人以上×各歳であっても、高齢層の一部を除き母集団一意にはならない（産業分類等の要素が入ると母集団一意の出現率は高まる）
 - ・攪乱的手法を導入することで匿名性の確保が可能
- 匿名データに関する提言
 - ・匿名化の度合いによって管理や貸与手続きの厳しさを変えることが必要
 - ・匿名化の度合いを高めパブリックユースファイルとして利用できるデータを作ることが望ましい

《調査研究の構成員》

座長 永瀬 伸子（お茶の水女子大学）
 川口 大司（東京大学）
 千田 浩司（NTTセキュアプラットフォーム研究所）
 南 和宏（統計数理研究所）

事務局 一橋大学経済研究所