

平成22年11月25日	資料1
第3回レセプト情報等の 提供に関する有識者会議	

総務省統計局における匿名データの提供について

平成22年11月25日

総務省 統計局 統計調査部 調査企画課

調査官 高田聖治

統計法における「匿名データ」とは

- 統計法の抜本改正（平成21年4月全面施行）により、新たに導入された制度
- 「一般の利用に供することを目的として調査票情報を特定の個人又は法人その他の団体の識別（他の情報との照合による識別を含む。）ができないように加工したもの」
(統計法第2条第12号)
- 匿名化のためには、直接的な識別情報（氏名、住所など）の削除だけでは不十分。
他の個体と識別が容易な特徴（例：大家族、超高齢者、超高額所得者）については、個人が特定できないよう、データの加工を行うことが必要。（「秘匿処理」）
- 学術研究の発展に資すると認められる場合等に、一般からの求めに応じ、匿名データを提供。
(統計法第36条)

匿名データの提供状況

- 現在、匿名データの提供を行っているのは、基幹統計調査（政府全体で52）のうち、統計局所管の4調査。
 - ・ 全国消費実態調査（5年周期調査、標本サイズ：約6万世帯）
 - ・ 社会生活基本調査（ “ ” 、標本サイズ：約8万世帯）
 - ・ 就業構造基本調査（ “ ” 、標本サイズ：約40万世帯）
 - ・ 住宅・土地統計調査（ “ ” 、標本サイズ：約350万住戸・世帯）

- 今後、いくつかの統計調査が追加される見込みであるが、
 - ・ 企業を対象とした統計調査は、匿名化が困難
（業種、企業規模(売上高、従業者数)だけで、有名企業は特定化されてしまう）
 - ・ 世帯対象の統計調査でも、どのような秘匿処理を行うかは、検討が必要

- 匿名データの作成にあたっては、統計委員会への諮問が必要（統計法第35条）
 - 匿名性が確保されていること、作成された匿名データの有用性等を審議

匿名データの作成方法

- ① 識別情報の削除：氏名、住所等の情報を削除
- ② リサンプリング：元の統計調査のレコード全てを匿名データ作成に用いるのではなく、再抽出（リサンプリング）したものをを用いる。（全消、社会調、就調：80%、住調：10%）
- ③ 裾切りによるレコード削除：特徴的な識別情報のあるレコードを削除。（「8人以上世帯」等）（アメリカ・センサス局では、0.5%ルールを採用）
- ④ トップ（ボトム）コーディング：極端に大きな（小さな）値は上限値を設けて頭打ちにする（年齢85歳以上を統合、等）
- ⑤ リコーディング：分類事項を粗く表示する、又は、連続値を階級値で表章（年齢を5歳階級で表章、地域区分を粗くする、等）

以下の手法は、諸外国では利用されているが、統計局では、現在、採用していない

- ・スワッピング：2つのレコード間で、一部の調査事項の値を入替え
- ・誤差の導入：一部の調査事項（年収、等）に誤差を導入

匿名データ提供に係る人的・技術的体制

- 仕組みの検討（匿名データ提供の方針の策定）

総務省統計局 調査企画課 二次利用推進係

- 提供の実務（利用者からの相談・申請受付、審査、データの複製・提供）

(独)統計センター 統計データ高度利用推進室 9名のうち2名※

※ オーダーメイド集計の提供も併せて担当

- 技術的研究（匿名化手法等の研究）

総務省統計局、(独)統計センター研究主幹、等

- ・ 連携協力機関（研究者からの相談を受け付け、申請を(独)統計センターに取次ぎ）

現在、一橋大、神戸大、法政大、情報・システム研究機構と協力協定を締結

匿名データ制度創設までの歩み

統計法の抜本改正以前：

調査票の統計上の目的外での使用は原則禁止

総務大臣の承認を得て使用の目的を公示した場合に、例外的に、使用可（「目的外使用制度」）

→行政機関における使用など

- ・ 1995年3月 統計審議会答申「統計行政の新中・長期構想」
標本データの提供について、おおむね2～3年を目途に専門的・技術的な検討を行う必要

これを受けて、統計局において、技術的・制度的な検討を行うとともに、以下の共同研究を実施。

- ・ 1996～98年度 一橋大学（社会科学統計情報研究センター）における研究
研究目的：法規面など社会的な問題、秘密保護の方法など統計技術上の問題 等
研究方法：利用者は集計プログラムを事務局に提出し、事務局でデータを集計
利用件数：58件
- ・ 2000～04年度 （財）統計情報研究開発センターにおける研究

研究目的：秘匿処理のユーザビリティの検証 等

研究方法：利用者を公募し、事務局が作成した匿名データを使って、事務局と共同研究

利用実績：44件

- ・ 2004～08年度 一橋大学（社会科学統計情報研究センター）における研究（「試行的提供」）

研究目的：匿名データ提供の実務に係る検討 等

研究方法：利用者を公募し、事務局が作成した匿名データを利用して研究

利用実績：132件

- ・ 2006年6月 統計制度改革検討委員会報告

統計制度の抜本的改革を提言

- ・ 2007年5月 統計法の全面改正

- ・ 2009年4月 改正統計法の全面施行

匿名データの提供制度の開始