

シミュレーション／分析における合成人口データベースの利用ルール

○村田忠彦 原田拓弥（関西大学）

Terms of Use in Simulations and Analysis Using Synthesized Populations Database

* T. Murata and T. Harada (Kansai University)

概要一 本稿では、実人口と実環境を対象としたリアルスケール社会シミュレーション (RSSS: Real-Scale Social Simulations) の開発を促進するために日本全国の合成人口を研究者向けに提供する際の、合成人口データの取り扱いにおける課題を整理し、個人情報の漏洩や推定につながらないようにするための要件を定める。

キーワード: 合成人口, 利用ルール, リアルスケール社会シミュレーション

1 はじめに

実人口と実環境を対象としたリアルスケール社会シミュレーション (RSSS: Real-Scale Social Simulations)¹⁾の開発にあたっては、まずシミュレーションの対象となる人口の合成が求められる。村田らは、日本における国勢調査をもとにして、日本全国の実人口と同じ規模の人口合成を行い、研究者向けの提供を開始している^{2,4)}。合成人口には、現時点では所得や就労している産業分類、将来的には、職業、教育歴などの個人情報が含まれる可能性があり、合成人口の提供が個人情報の漏洩や特定につながらないようにするため、合成人口データベースの利用ルールを定める。

2 合成人口の配布と利用ルール

RSSSの研究を進める際、各研究者は対象の地域の人口の合成を行なうが、シミュレーションに必要な属性しか考慮せず人口合成が行われるため、使用しない属性に関する実統計との整合性は考慮されないのが普通である。個々のRSSSが異なる人口構成を用いているのは、将来的に各シミュレーションを接合し、より大規模で複合的なシミュレーションを構築する際の問題となる。そこで、合成の際に利用した統計を明示し、RSSSに利用可能な合成人口が統一的に用いられることが望ましい。Table 1に示すように、合成人口の提供を行っているのは、アメリカ、イギリス、日本である。

Table 1に示す米英の各公開母体は研究者からの申請に応じて、合成人口データの提供を行なっているが、その利用方法については特に明記されていない。RTI Internationalは、合成人口ビューワー利用者に対して提示されるデータが個人に紐づかないことを示しているのみである。

本プロジェクトでは、以下のような要件で合成人口

データの提供を行う。

- 1) 合成人口は、現実の世帯や個人の情報を含まない。
- 2) 合成人口は、世帯や個人の現実の統計と同じ特徴をもつ。
- 3) 合成人口は、合成時に使用していない統計との整合性は保証しない。
- 4) 合成人口は、統計が新たに公開された際に更新されることがある。
- 5) 合成人口を用いたシミュレーションや分析は、複数の合成人口データを用いて行わなければならない。
- 6) 合成人口を用いたシミュレーションや分析の結果は、現実の世帯や個人に結びつくような形で公開してはならない。

上記要件1と2にあるように、合成人口は実際の世帯や個人の情報を含まない。しかし、データを公開された世帯や個人にとっては、地図上の自宅に紐づけられた情報で、個人のプライバシーが侵害されたように感じるため、要件5に示したように、複数の合成人口データを用いて、シミュレーションや分析を行うように定めると共に、要件6に示すように、現実の世帯や個人に直接結びつくような結果の開示を禁止する。

3 おわりに

本稿では、合成人口提供の際の要件をまとめた。特に利用者によるシミュレーションや分析の結果の開示において、実際の世帯や個人の特定につながらない情報の開示が必要である。

謝辞 本研究の一部は、JSPS 科研費 17K03669, 科学技術融合振興財団, 立石科学技術振興財団の助成を受け、大阪大学サイバーメディアセンターの大規模計算機システムを利用して得られたものです。

参考文献

- 1) 村田, 原田, 杉浦: リアルスケール社会シミュレーションの実現に向けて, 日本シミュレーション学会論文誌, **34-4**, 58/62 (2017)
- 2) T. Murata, T. Harada, D. Masui: Comparing Transition Procedures in Modified Simulated-Annealing-Based Synthetic Reconstruction Method Without Samples, *SICE JCMSI*, **10-6**, 513/519 (2017)
- 3) T. Harada, T. Murata: Projecting Households of Synthetic Population on Buildings Using Fundamental Geospatial Data, *SICE JCMSI*, **10-6**, 505/512 (2017)
- 4) 杉浦, 村田, 原田: 賃金構造基本統計調査に基づく合成世帯集団の労働者への所得の割当て, システム制御情報学会論文誌, **32-2**, 70/79 (2019)

Table 1: 人口合成の公開例

公開母体	合成対象	使用統計
RTI International, USA	アメリカ合衆国全州 人口: 3 億人	2010 US Decennial Census 2007-2011 American Community Survey
CDRC: Consumer Data Research Center, UK	England & Wales, UK 人口: (53 + 3) 百万人	2011 UK Census
関西大学, 日本	日本全都道府県 人口: 1 億 2 千万人	2000 年国勢調査 2005 年国勢調査 2010 年国勢調査 2015 年国勢調査