

携帯電話事業者の提供する GPS に基づく人口統計の信頼性に関する一考察

東京都市大学大学院（現 株式会社建設技術研究所） 正 会 員 ○上原 涉豊
 法政大学 正 会 員 今井 龍一
 株式会社サーベイリサーチセンター 正 会 員 高野 精久
 KDDI株式会社 非 会 員 若井 幸夫

1. はじめに

我が国の各都市では、人口減少および超高齢化社会に相応な施策に取り組む必要がある。その中で定住人口に限らない交流人口の把握は、重要な1項目である。一部の地域では、概ね10年に1度実施されるパーソントリップ調査で属性、移動目的や移動手段を含めた人々の移動実態を把握可能である。しかし、調査間隔が長い場合、データの鮮度に課題がある。

ICTの技術動向に着目すると、情報通信機器は年々普及が進み、2018年時点で日本国民の96%が携帯電話を保有している¹⁾。スマートフォンに限っても国民の79%が保有している¹⁾。近年、携帯電話基地局の運用データから生成された人口動態統計やスマートフォンアプリケーションのGPS情報から生成された流動人口データの流通が進んでいる。人口動態統計は、統計として確立されており、街づくりで使われている。一方、GPS情報から生成された人口統計のデータセットは、携帯電話の契約ユーザを対象にしているため、十分なサンプル数を元に生成されているが、人口動態統計同様に統計的な信頼性の検証が必要である。

本研究の目的は、携帯電話から収集したGPS情報を元に生成した人口統計の統計的な信頼性の考察とした。

2. 本研究で検証するデータの概要

本研究で精度検証するデータは、KDDI株式会社が提供するKDDI Location Data(以下、「KLD人口」とする。)である。このデータは、同社のスマートフォンユーザのうち承諾を得た数百万ユーザを対象としている。対象ユーザから定期的にGPS位置情報を取得し、契約情報の属性(性別・年代)と組み合わせ、メッシュごとに拡大推計処理を施して集計している。したがって、人口動態統計と同様、個人情報・プライバシーが保護されたオプトインデータである。主な収録項目を表-1に示す。

滞在と移動の判定は、同一メッシュに15分以上滞在した場合を滞在、15分未満の滞在を移動と判定している。メッシュは、4次(500m四方)～6次(125m四方)で集計されている。

3. 評価方法

本研究では、KLD人口と悉皆調査である国勢調査の人口等の基本集計に関する事項(以下、「国調人口」とする。)との比較で統計的な信頼性を評価する。KLD人口は、1時間単位で集計されているため、人々の動きが少なく居住人口と考えられる3時台に滞在と判定されたデータを用いる。対象年月は、2019年6月の平日平均および休日平均とした。国調人口は、現時点で公開されているデータで最新の2015年に調査された20歳以上人口を用いる。

比較の指標には、大藪ら²⁾が用いている偏差率を用いる。偏差率は、KLD人口からKLD人口と国調人口の平均(以下、「平均人口」とする。)を引いた値を平均人口で割った指標である。人口規模に係わらず両者の差を-1～1で評価できる。偏差率は、-1に近いほど国調人口が多く、1に近いほどKLD人口が多いことを表し、0が両者の差が最も小さい状態であることを表す。KLD人口を K_i 、国調人口を C_i としたとき、偏差率は、式-1で表せる。

4. 評価結果

本研究では、東京都全域を対象に市区町村および4次メッシュごとに偏差率を算出した。市区町村ごとの偏差率は、

表-1 収録項目

項目	内容
年月	平日、休日別
時間帯	1時間単位
メッシュコード	500m/250m/125m単位の標準地域メッシュコード
市区町村コード	5桁
年代	20歳代、30歳代、40歳代、50歳代、60歳以上の5区分
性別	男女別
平均滞在人口	平日、休日別の平均人数
平均移動人口	平日、休日別の平均人数

$$\text{偏差率} = \frac{K_i - (K_i + C_i)/2}{(K_i + C_i)/2} \dots (1)$$

キーワード 携帯電話, GPS, 位置情報, 人口推定

連絡先 〒162-0843 東京都新宿区市谷田町2-33 法政大学 TEL: 03-5228-1448

島嶼部以外では、ほぼ違いはみられなかった。市区町村別の平均偏差率は-0.20 であり、やや国調人口の方が大きかった。次に、4 次メッシュごとの算出結果を図-1 に示す。図-1 より、都心部である東京駅、新宿駅および品川駅周辺では、国調人口に比べて KLD 人口が非常に多くなっていることが読み取れる(図-2 参照)。また、臨海副都心では、国調人口のデータがなく評価できないメッシュもあったが、KLD 人口をみると、人口は多いエリアであった。一方、多摩西部では、KLD 人口の秘匿処理の影響が大きく評価できないメッシュが多数みられた(図-3 参照)。次に、偏差率を 20 等分し、それぞれのメッシュ数を集計した結果を図-4 に示す。図-4 より、平休日で比較した場合、性別で比較した場合のいずれにおいても大きな違いはみられなかった。どの条件でも偏差率は、-0.2 前後にピークがあり、KLD 人口に比べて国調人口の方が多いメッシュが多くなっていた。また、両者の相関関係を図-5 に示す。図-5 より、属性によらず決定係数 R^2 は 0.8 程度であることが分かった。

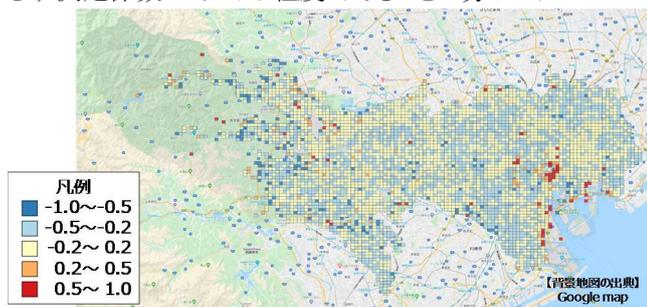


図-1 東京都全域のメッシュごとの偏差率 (性別による区分なし)



図-2 東京駅・新宿駅・品川駅周辺の偏差率

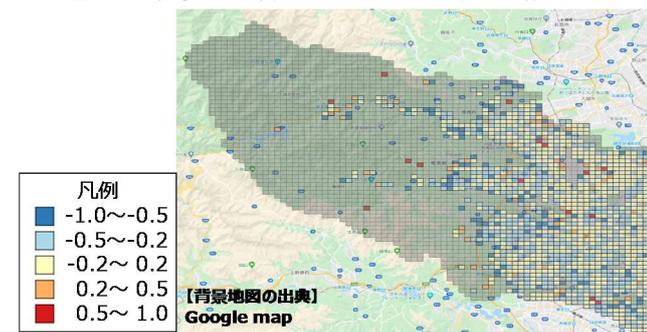


図-3 東京都多摩西部の偏差率

5. まとめ

本研究では、KDDI 株式会社の携帯電話契約ユーザから収集した GPS 位置情報で生成した移動滞在データの概要を調査した。KLD 人口の統計的な信頼性は、KLD 人口の夜間人口と国調人口とを比較して検証した。その結果、夜間人口では、曜日や性別に係わらず、やや国調人口の方が多かったが、KLD 人口と国調人口の相関関係は高く、代表性を有していると考えられる。今後の課題として、他の空間解像度や携帯電話基地局の運用データ等の他のデータとの比較が挙げられる。

謝辞: 本研究の遂行にあたり、日本工営株式会社の金井翔哉氏、株式会社福山コンサルタントの国分恒彰氏、中央復建コンサルタンツ株式会社の松島敏和氏、一般財団法人計量計画研究所の矢部努氏から貴重な意見を賜った。ここに記して感謝の意を表する。

参考文献

- 1) 総務省：平成 30 年通信利用動向調査、<<https://www.soumu.go.jp/johotsusintokei/statistics/statistics05.html>>, (2020.2.22 閲覧)
- 2) 大藪勇輝, 寺田雅之, 山口高康, 岩澤俊弥, 萩原淳一郎, 小泉大輔：モバイル空間統計の信頼性評価, NTT DOCOMO テクニカル・ジャーナル, Vol.20, No.3, pp.17-23, 2012.

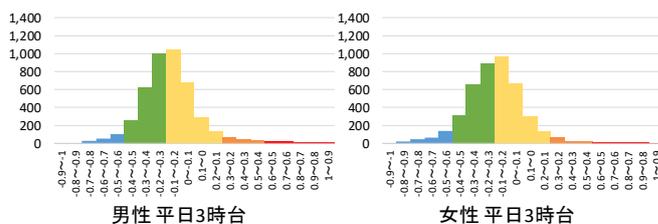
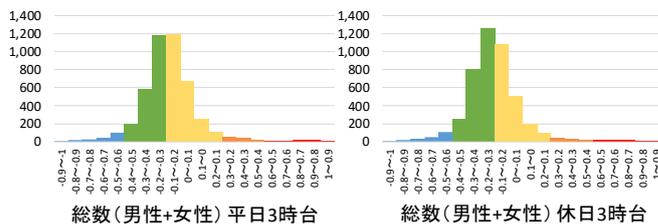


図-4 属性および平休日別の偏差率の分布

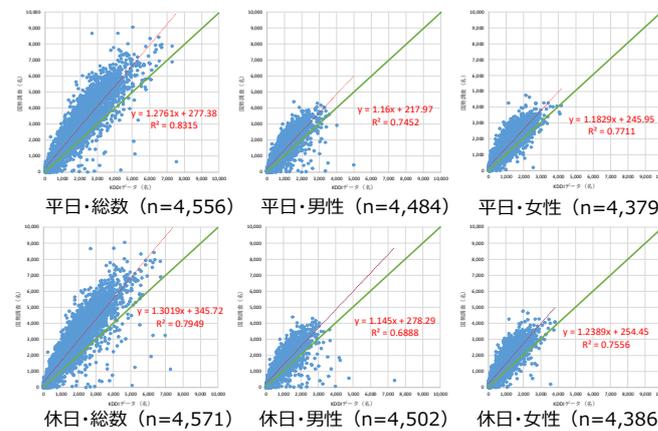


図-5 KLD 人口と国勢調査人口の相関