

ブートストラップ法を用いた OD 表の区間推定に関する基礎研究

熊本大学大学院

学生会員

○河岡英明

熊本大学

正会員

円山琢也

1. はじめに

都市圏レベルの交通実態を把握するために、我が国では、道路交通センサス、パーソントリップ調査（以下、PT 調査）を代表とする大規模な交通調査がなされてきた。わが国では OD 表の精度を確保するため、大規模なサンプル調査が実施されている。PT 調査は、あらゆる交通機関を用いたトリップを同時に調査していることにより交通機関別分担の解析ができること、トリップの起点から終点までを調べ、これを交通の主体である人の特性や起終点の都市活動や土地利用と結びつけて解析できる点などで優れた交通調査法である。

しかしながら、一般に交通調査の結果には様々な誤差が含まれる。例えば、1. 標本抽出時における偶然性に基づく標本誤差、2. 記入漏れなどの回答誤差、3. 調査未回答者による誤差、4. データ整理時の誤差があり、調査結果の精度評価が求められる。本研究では、上記のうち 1 のみを対象として、交通調査により得られたデータの信頼区間を求めることを目指す。より具体的には、ブートストラップ法を用いて PT 調査 OD 表の区間推定を実施することを目的とする。

2. 方法の概要

本研究は、2012 年熊本都市圏パーソントリップ調査（以下、熊本 PT 調査）を対象とした。OD 表作成までの概要を以下に示す。

(1) 利用したデータ

OD 表作成のため、熊本 PT 調査のマスターデータを用いた。調査の対象地域は熊本都市圏（熊本市を中心とした 5 市 6 町 1 村）で、総人口（行政区域人口）は約 104 万人である。このデータを集計し、実際の熊本都市圏の人口に拡大するため、熊本都市圏を 50 個の B ゾーンに分け、ゾーンごとの性・年齢別人口のデータも用いた。本研究では、都市圏外との出入りを除く熊本都市圏内のみの移動を対象を限定した。

(2) 標本抽出

本研究は、OD 表を作成する際に、ブートストラップ法による無作為標本抽出を行った。熊本 PT 調査のマスターデータから熊本圏内を移動するトリップを持つ個人データ 73,266 個を用いた。その中から 10,000 個の重複を許す無作為標本抽出を行う。抽出率は 13.6%である。

(3) データの拡大

無作為抽出した個人データの中で、性別・年齢(5 歳単位)・居住地が一致するものを集計し、ゾーンごとの性・年齢別人口に拡大した。

(4) OD 表作成

地域内の交通流動パターンを面的にとらえることを目的として、標本抽出を行ったデータを集計し OD 表とする。(2)から(3)を 100 回繰り返す、熊本都市圏全体の OD 表の分布を作成する。

(5) 区間推定方法

OD 表を作成してから、熊本都市圏内で最も移動の多い熊本市内の OD 表の区間推定を行う。区間推定の方法として、ブートストラップ分布から直接信頼区間を得るパーセンタイル法を用いる。パーセンタイル法では、無作為標本 y_1, \dots, y_n から、復元抽出により互いに独立なブートストラップ標本 Y^*_1, \dots, Y^*_n を抽出し、ブートストラップ推定量 $\hat{\theta}^* = \hat{\theta}(Y^*_1, \dots, Y^*_n)$ の値を計算する。これを 100 回繰り返す $\hat{\theta}^*_1, \dots, \hat{\theta}^*_B$ を求め、小さい順に並び替える。 $0 < p < 1$ に対して pB 番目に小さい数値を $\hat{\theta}^*_{(p)}$ と表し、 pB が整数値でなければ、となりあう値を線形補間する。 $p = \alpha/2$ および $p = 1 - \alpha/2$ とおくと、 $100(1 - \alpha)\%$ 両側信頼区間を次式で与える¹⁾。

$$[\hat{\theta}^*_{(\alpha/2)}, \hat{\theta}^*_{(1-\alpha/2)}] \quad (1)$$

とくに θ の 95%信頼区間は(2)で表される。

$$[\hat{\theta}^*_{(0.025)}, \hat{\theta}^*_{(0.975)}] \quad (2)$$

3. 結果と考察

(1) 人口と総トリップ数

熊本市内の東区、西区、南区、北区の各ゾーンを出発地とし、中央区を到着地とする OD 交通量の分布を求めた。各区の総人口を表 1 に示す。また、各区から中央区に移動するトリップ数の合計値の平均値を図 1 に示す。図 1 から中央区に向かうトリップで一番多いのは東区であることが分かり、北区が少ないことが分かった。東区は中央区に次いで人口が多いためトリップも多くなったと考えることが出来る。しかし、北区は西区、南区より人口が多いにもかかわらずトリップ数が最も少なかった。北区の場合については、職場や年齢層などの個人属性によって中央区へのトリップ数が少なく

表 1 2012 年熊本市人口推移 (人)

	総人口	男性	女性
中央区	196,521	89,944	106,577
東区	190,482	90,442	100,040
西区	93,620	43,709	49,911
南区	125,458	59,419	66,039
北区	147,595	70,433	77,162

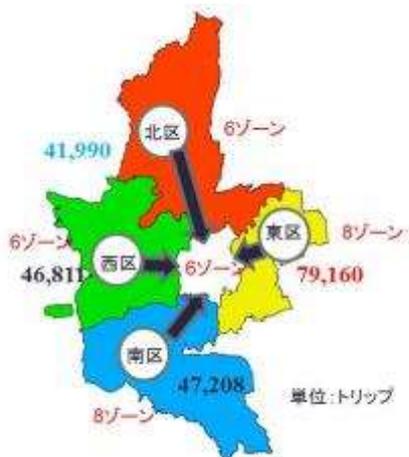


図 1 中央区へのトリップ数の平均値

なつたと考えられる。

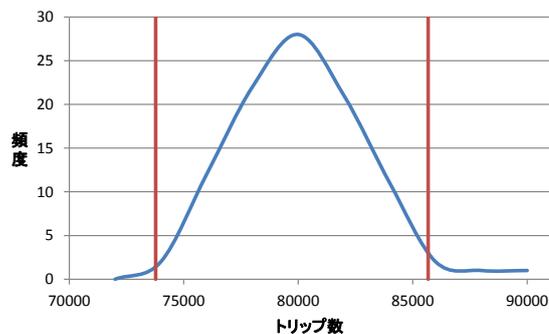
(2) 区間推定

ブートストラップ法によって得られた 100 個の OD 表からトリップ数についてヒストグラムを作成し、パーセントイル法で 95%信頼区間(図中の赤線)を求めたものを図 2(a)から(d)に示す。東区、北区ではトリップ数にバラつきが少ないことが分かり、西区、南区ではトリップ数にバラつきがあることが分かった。各区のトリップの分布には、個人属性が大きく関わっていると考えられる。

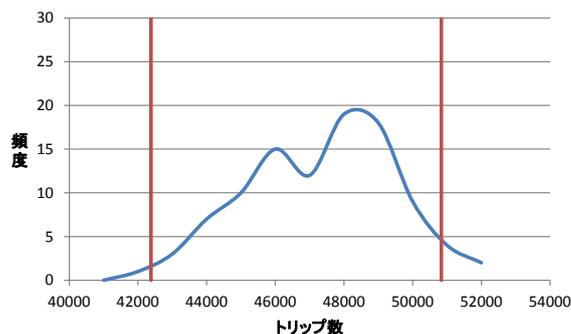
4. 今後の展望

各区のからのトリップの平均値や信頼区間に違いは見られたが、その要因として各区の個人属性が大きく関係しているのではないかと考えられるため、区別に基礎分析を細かく行い、トリップ数にどのような影響があるのか、調べる必要がある。

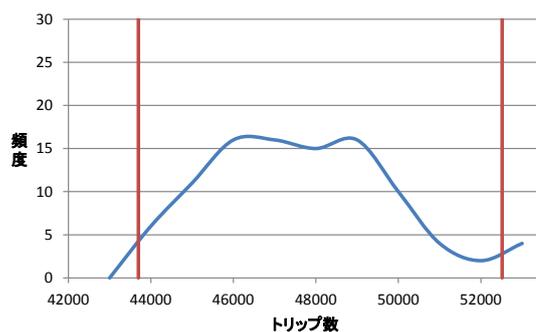
本研究の結果を利用して、配分リンク交通量や利用者便益の区間推定を今後実施したい。また、調査参加選択モデルを組むことによって²⁾、調査未回答者による誤差を考慮した区間推定も可能と考えられる。



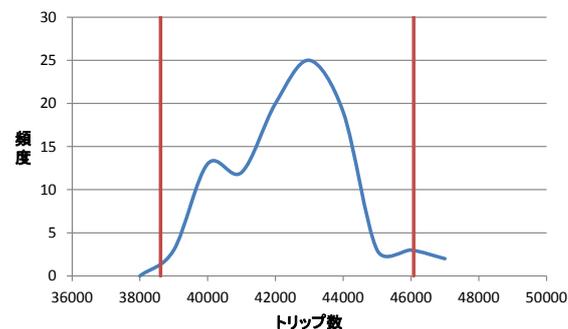
(a) 東区→中央区



(b) 西区→中央区



(c) 南区→中央区



(d) 北区→中央区

図 2 中央区への区別のトリップ平均値の分布

参考文献

- 1) 下平英寿: 「21 世紀の統計科学 3 数理・計算の統計科学」, 統計計算の展開と統計科学 第 III 部, pp.193-199, 2011.
- 2) 円山琢也, 平原雄太郎: スマホ調査を実施した PT 調査における調査参加選択行動モデルの構築, 第 49 回土木計画学研究発表会, 2014.