

交通調査におけるトリップ出発時刻の丸め誤差発生確率の推定

熊本大学 学生会員 ○佐藤嘉洋 上原一輝 正会員 円山琢也

1. 研究の背景と目的

交通調査に限らず多くの社会調査では、回答に少なからず「誤差」が含まれていることが一般的に知られている。交通調査の分野での非標本誤差の一つとして、トリップ発生時刻の丸め誤差が挙げられ、多くの研究でこの誤差について指摘が行われている(例えば、Levinson and Kumar¹⁾など)。パーソントリップ調査では、回答された出発時刻および到着時刻をもとにトリップ所要時間を計算し、交通手段選択モデル等の推定を行うなど、調査参加者が行動した時刻は大きな役割を担う。しかし、回答された時刻がどのように丸められて回答されているのか、丸め誤差の発生確率に関する検討は十分とは言えない。そこで本研究では、熊本都市圏パーソントリップ調査(以下、熊本PT調査)を事例に、時刻の丸め誤差の傾向を確認し、その発生確率について推定を行うことを目的とする。

2. 先行研究と使用するデータ

本研究に類似する先行研究として、Rietveld²⁾はオランダの全国交通調査のデータをもとに丸め誤差の発生確率の推定を行っている。本研究での推定はRietveldの手法を用い、推定結果の比較も目的としている。使用するデータは先述のとおり2012年に実施された熊本PT調査で、分析対象は出発時刻(以下、補足がない場合は出発時刻の「分」を単に出発時刻と表す)が不明のデータを除外した240,006トリップである。なお、トリップの到着時刻については、時刻が予め決まっているイベントを目安として回答される傾向があるため、本研究では出発時刻の丸め誤差発生確率のみに着目する。

3. 報告された出発時刻の分布

図1に出発時刻の分布を示す。多くの先行研究が指摘しているように、5分単位での丸め誤差が発生していることが見て取れるが、Rietveldの結果と大きく違う点は10分と15分単位の丸め誤差の関係である。Rietveldの研究、つまりオランダの調査では15分単位で出発するトリップが突出しており、アメリカの調査でも同様の傾向が見られると報告されている。一方、熊本PT調査では15分単位で出発するトリップよりも10分単位で出発するトリップの方が多い。この背景には、欧米諸

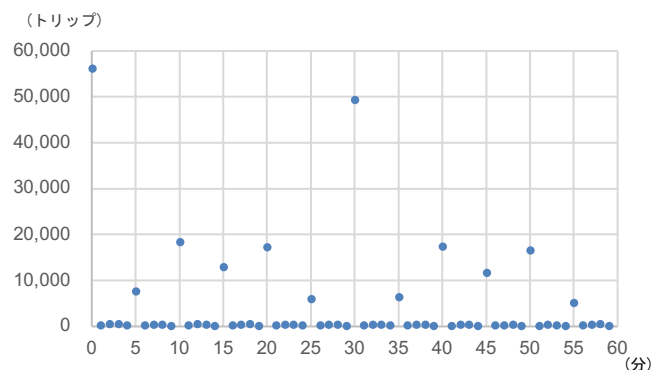


図1 出発時刻の分布

国では4分の1時間、つまり“Quarter-hour”の概念が強い一方、日本では10分単位で捉える方がより自然である、という考え方の違いが影響している可能性がある。

4. 丸め誤差の発生確率推定

4.1 発生確率の定義

回答者の真の出発時刻が m 分のとき、5分単位での丸め誤差の発生確率を $p_{m,5}$ とすると、定数 a_5, b_5 を用いて丸め誤差発生確率を以下のように定義する。

$$p_{m,5} = a_5 + b_5 \cdot d_{m,5} \quad d_{m,5} = 1, 2 \quad (1)$$

ここで、 $d_{m,5}$ は m から最も近い5分単位の時刻までの差を表し、例えば $m=13$ のときは15分が最も近い5分単位の時刻のため、 $d_{13,5}=2$ である。他に丸め誤差が発生する可能性としては10分単位、15分単位、30分単位、60分単位が考えられ、これらを丸められうる値と定義する。これらについても式(1)と同様にそれぞれ以下のように定義する。なお、3.での分析を鑑み、Rietveldの手法では用いられていない $p_{m,10}$ を新たに追加した。

$$p_{m,10} = a_{10} + b_{10} \cdot d_{m,10} \quad d_{m,10} = 1, 2, \dots, 5 \quad (2)$$

$$p_{m,15} = a_{15} + b_{15} \cdot d_{m,15} \quad d_{m,15} = 1, 2, \dots, 7 \quad (3)$$

$$p_{m,30} = a_{30} + b_{30} \cdot d_{m,30} \quad d_{m,30} = 1, 2, \dots, 15 \quad (4)$$

$$p_{m,60} = a_{60} + b_{60} \cdot d_{m,60} \quad d_{m,60} = 1, 2, \dots, 30 \quad (5)$$

4.2 最尤法としての定式化

回答者が真に m 分に出発した確率を g_m をとくと、 m 分にもっとも近い5分単位の時刻に丸められる確率は、丸め誤差発生確率を用いて $g_m \cdot p_{m,5}$ と表すことができる。10分単位から60分単位に丸められる確率に関しても、式(2)~(5)を用いて同様に記述できる。ここで、回答者が m 分と回答する確率を q_m とおくと、 m が丸め

られうる値でないときは以下のように記述できる.

$$q_m = g_m \cdot [1 - p_{m,5} - p_{m,10} - p_{m,15} - p_{m,30} - p_{m,60}] \quad (6)$$

式(6)は, 調査参加者が真に m 分に出発した確率と, どの値にも丸められない確率の積を表している.

一方, m が丸められうる値であるとき, 例えば $m=15$ のときは, 前後の値から丸められる確率を加えて以下のように表される. なお, $m=15$ のとき $p_{m,5} = p_{m,15} = 0$ であるため, 式(7)の第3項は式(6)の右辺と等価である.

$$\begin{aligned} q_{15} = & [g_8 \cdot p_{8,15} + \dots + g_{22} \cdot p_{22,15}] \\ & + [g_{13} \cdot p_{13,5} + \dots + g_{17} \cdot p_{17,5}] \\ & + g_{15} \cdot [1 - p_{15,10} - p_{15,30} - p_{15,60}] \end{aligned} \quad (7)$$

ここで, 真に出発した時刻は一様に分布しているとの仮定を置き, $g_m = 1/60$ とする. このとき, N_m を出発時刻が m 分と回答されたトリップの総数とすると, 同時生起確率 L は以下のように表される.

$$L = \prod_{m=0}^{59} q_m^{N_m} \quad (8)$$

対数尤度関数は以下のとおりである.

$$\ln L = N_0 \ln q_0 + N_1 \ln q_1 + \dots + N_{59} \ln q_{59} \quad (9)$$

ただし, q_m およびその中に含まれる $p_{m,5}, \dots, p_{m,60}$ は確率変数であるため, 以下の条件が必要となる.

$$0 \leq q_m, p_{m,5}, p_{m,10}, p_{m,15}, p_{m,30}, p_{m,60} \leq 1 \quad (10)$$

以上より, 丸め誤差発生確率の推定は式(10)を満たしながら式(9)の尤度を最大化する, 制約条件つき尤度最大化問題として定式化できる.

5. 推定結果

推定結果を表 1 に示す. 最も近い丸められうる値との差に影響する $b_5, \dots, b_{60}$ については, すべて負の値をとった. これは出発時刻が丸められうる値まで離れているほど丸め誤差の発生確率が低くなることを示している. ただし, 5 分単位以外はいずれも 1 分あたり 1% 未満で, 影響は限定的である. 丸められうる値との差に影響しない $a_5, \dots, a_{60}$ では, 5 分単位の値 (a_5) が最も大きい一方, 15 分単位の値 (a_{15}) が 5 分単位だけでなく 10 分単位, 30 分単位と比べても小さい. これは丸められうる値との差の大小に関わらず, 15 分単位での丸め誤差発生確率が低いことを意味しており, Rietveld の結果とは大きく違う傾向を示している点である.

具体例として, 表 2 に 10:19 に真に出発したときの丸め誤差発生確率を整理した. 表より, およそ 63% の確率で最も近い 5 分単位/10 分単位の 10:20 に丸められる可

表1 丸め誤差発生確率モデル推定結果

係数	最尤推定値	標準誤差	係数	最尤推定値	標準誤差
a_5	0.452	0.0027	b_5	-0.0575	0.0011
a_{10}	0.240	0.0023	b_{10}	-0.0059	0.0004
a_{15}	0.092	0.0018	b_{15}	-0.0032	0.0002
a_{30}	0.243	0.0022	b_{30}	-0.0011	0.0001
a_{60}	0.028	0.0013	b_{60}	-0.0002	0.7e-06

表2 10:19に真に出発したときの丸め誤差発生確率

丸め誤差	60分	15分	なし	5分/10分	30分
回答時刻	10:00	10:15	10:19	10:20	10:30
発生確率	2.4%	7.9%	3.7%	62.9%	23.1%

能性があり, 丸められず回答される確率はわずか4%程度に過ぎない. 表1および表2から, 丸め誤差は5分単位で最も多く発生し, 15単位, 60分単位での丸め誤差発生および丸められず回答される確率は低いことがわかる.

6. おわりに

本研究では, 熊本 PT 調査でのトリップ出発時刻の丸め誤差発生確率を推定し, 誤差の発生に関する傾向を明らかにした. 具体的な成果を以下に示す.

- (1) 丸め誤差は 5 分単位での発生確率が最も高く, およそ 35~40% の確率で発生すると推定される. また, 真の出発時刻が丸められうる値から離れているほど 5 分単位の丸め誤差発生確率は低くなる.
- (2) 欧米での研究と比較して, 熊本 PT 調査では 15 分単位での丸め誤差発生確率は低い傾向にある.

なお, PT 調査では本人以外が代理で回答していることも考えられる. 今後の展望としては, 集団意思決定モデルを応用した代理回答確率³⁾を用いて丸め誤差発生確率を推定し, 代理回答が丸め誤差に与える影響について評価を行いたい.

【参考文献】

- 1) Levinson D. M., Kumar A: A Multi-Modal Trip Distribution Model: Structure and Application, Transportation Research Record, Vol.1466, pp.124-131, 1994.
- 2) Rietveld P: Rounding of Arrival and Departure Times in Travel Surveys: An Interpretation in Terms of Scheduled Activities. Journal of Transportation Statistics, Volume 5, Issue 1, pp. 71-82, 2001.
- 3) 細谷謙太, 川野倫輝, 渡邊萌, 佐藤嘉洋, 円山琢也: 集団意思決定を考慮した世帯単位の交通調査回答行動分析, 第 57 回土木計画学研究発表会(春大会), 2018. 06.