

交通調査のトリップ出発時刻の丸め誤差 ：都市間比較

佐藤 嘉洋¹・円山 琢也²

¹学生会員 熊本大学大学院自然科学教育部工学専攻（〒860-8555 熊本市中央区黒髪2-39-1）

E-mail:yo-sato@kumamoto-u.ac.jp

²正会員 熊本大学准教授 くまもと水循環・減災研究教育センター（〒860-8555 熊本市中央区黒髪2-39-1）

E-mail: takumaru@kumamoto-u.ac.jp

パーソントリップ調査などの交通調査では、トリップの出発・到着時刻を繰り上げまたは繰り下げて回答する「丸め誤差」が発生することが知られている。本研究では、独立行政法人国際協力機構が実施した海外19都市の調査をもとに、トリップ出発時刻の丸め誤差発生傾向を比較した。丸め誤差発生確率モデル推定の結果、60分単位での回答の割合が特に高い、15分単位より10分単位の回答が多い、00分単位の回答も多いが5,10,15分単位での回答も見られる、60分単位での回答は比較的少なく、15分単位での回答の傾向が見られる、の4つに類型化できることが明らかになった。また、旅行時間の分布と丸め誤差発生傾向を比較し、60分単位の丸め誤差が起きにくい都市は移動が比較的短時間となる傾向などを示した。

Key Words : rounding error, travel survey, departure time, intercity comparison, travel time

1. 研究の背景と目的

一般的な社会調査において、得られた回答には少なからず誤差が存在しており、その誤差が結果に与える影響が問題となっている。社会調査の回答に含まれる誤差は標本誤差と非標本誤差に分けられるが、Tourangeau¹⁾が主張しているように近年では非標本誤差に関心が集まってきた。非標本誤差とは標本の抽出方法によらず、調査客体の行動等が原因となって起きる誤差のことであり、無回答や回答に含みうる誤差も非標本誤差と言える。パーソントリップ調査をはじめとする交通調査においては、トリップ出発、到着時刻(以下、回答時刻とする)²⁾または移動距離³⁾に非標本誤差が見られることが多く、実際の時刻または距離をある値に繰り上げまたは繰り下げ、切りの良い数字として回答する「丸め誤差」が起きやすいことが知られている⁴⁾。特に回答時刻においては、旅行時間の算定に関連するため、これら回答時刻の丸め誤差がその後の分析において交通手段選択モデルや時間帯別交通需要分析などに与える影響が懸念される⁵⁾。したがって、回答時刻の信頼性に関する精査は重要な課題と言える。

先行研究では、熊本都市圏パーソントリップ調査の回答時刻の傾向を分析し⁶⁾、Rietveld⁷⁾が提案した丸め誤差発生確率モデルを用いて欧米の調査と回答の傾向が異なる

ことなどの知見を得ている¹⁰⁾。丸め誤差の傾向が国によって異なるのは興味深い知見であるが、より幅広い国際比較は有用であろう。ただし、このような研究例は筆者の知る限り存在しない。そこで本研究では、

- ・トリップ出発時刻に関する丸め誤差発生確率を推定し、都市間の比較を行うこと
- ・推定された丸め誤差発生確率から、都市によるトリップ出発時刻の回答傾向を類型化すること
- ・旅行時間とトリップ出発時刻の回答傾向との関係を明らかにすること

を通して、トリップ出発時刻で起きる丸め誤差に関して都市による違いを明らかにすることを目的とする。

本稿の構成は以下のとおりである。まず、2章で使用データの説明および先行研究で用いた丸め誤差発生確率モデルの紹介を行い、都市ごとにパラメータの推定を行う。続いて3章では、2章で推定された結果からクラスタ分析による類型化を行い、トリップ出発時刻の分布と比較することで類型ごとに回答時刻の傾向の違いを明らかにしていく。さらに4章では旅行時間と回答時刻の傾向を比較し、最後に5章で本研究のまとめと今後の展望について整理を行う。



図-1 分析対象都市の位置図

表-1 都市交通分野の開発調査における交通調査統合データ概要 (独立行政法人国際協力機構作成)

No	都市名 (国名)	調査実施年	本研究におけるサンプルサイズ
1	Manila(フィリピン)	1996	471,035
2	Chengdu(中国)	2000	67,961
3	Managua(ニカラグア)	1998	54,137
4	Belem(ブラジル)	2000	59,529
5	Bucharest(ルーマニア)	1998	143,311
6	Cairo(エジプト)	2001	268,360
7	Kuala Lumpur(マレーシア)	1999	218,460
8	Ho Chi Minh(ベトナム)	2003	262,372
9	Hanoi(ベトナム)	2005	188,949
10	Nairobi(ケニア)	2005	38,484
11	Lima(ペルー)	2003	270,357
12	Dar es Salaam(タンザニア)	2007	26,775
13	Da Nang(ベトナム)	2008	50,509
14	Ulaanbaatar(モンゴル)	2008	38,015
15	Dhaka(バングラデシュ)	2009	153,654
16	Nairobi(ケニア)	2013	37,777
17	Maputo(モザンビーク)	2012	52,668
18	Phnom Penh(カンボジア)	2012	86,884
19	Karachi(パキスタン)	2011	217,875

2. 使用データおよび丸め誤差発生確率推定モデルの推定

(1) 使用するデータ

本研究では、都市交通分野の開発調査における交通調査統合データ (独立行政法人国際協力機構作成) を使用する。このデータを利用した先駆的な研究例として、Hyodo *et al.*¹¹⁾では主にアジア、中南米の13都市で比較を行い、トリップ出発時の分布や生成原単位、手段分担率などに関する都市間の違いを明らかにしている。本研究の分析対象とした都市の概要は図-1、表-1に示す。1990年代後半から2010年代前半にかけて、主にアジア、アフリカ、中南米の都市で実施された調査データに基づき分析

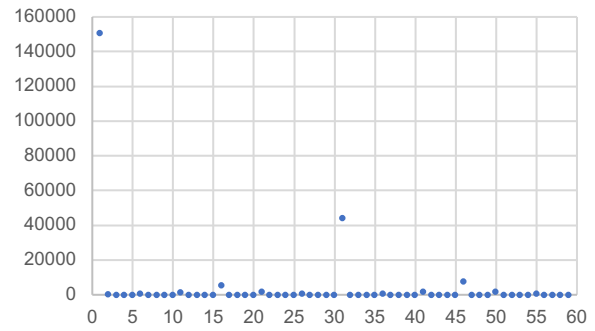


図-2 トリップ出発時刻の分布 (Karachi, 2011)

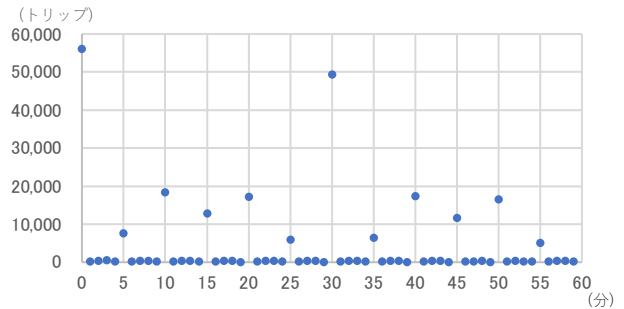


図-3 トリップ出発時刻の分布 (熊本PT調査)^{8),10)}

を行う。なお、本研究の対象は、マスターデータにおいて外出していないトリップおよびトリップ出発時刻が不明のトリップを除いたトリップ数をサンプルサイズとし、拡大係数は使用しない。分析の概要として、対象都市のうち、Karachi (2011)を例として図-2にトリップ出発時刻の分布を示す。参考に、図-3には先行研究^{8),10)}における熊本都市圏パーソントリップ (PT) 調査¹²⁾の出発時刻の分布を掲載するが、熊本PTと比較してもKarachiのデータは00分に出発したと回答したトリップのが多数を占める。他の都市においても概ね同様の傾向が見られ、本データ群の特徴と言える。このように偏りが大きいデータにおいても、丸め誤差発生確率モデルの推定が丸め誤差発生に関する国際都市間比較に有用であるか確認ができる点で本研究の意義は高い。

(3) 丸め誤差発生確率の定式化

ここでは、丸め誤差発生確率モデルの定式化を行う。ベースとなるモデルはRietveld⁹⁾、本研究で用いるモデルの詳細は先行研究¹⁰⁾を参照されたい。先行研究と同様に、60分単位(00分)、30分単位、10分単位(10分、…、50分)、15分単位 (15分、…、45分)、5分単位 (5分、25分、…、55分) を本研究では「丸められうる値」と定義する。

回答者の実際の出発時刻が 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、パラメータ 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。詳細は割愛するが、他の丸められうる値について、使用データ群における丸められない値の回答傾向を鑑み、先行研究と同様に最も近い丸められうる値までの差によらず一定であるとした。つまり、次の式(2)~式(5)で表される。

$$p_{m,10} = a_{10} \quad (2)$$

$$p_{m,15} = a_{15} \quad (3)$$

$$p_{m,30} = a_{30} \quad (4)$$

$$p_{m,60} = a_{60} \quad (5)$$

$p_{m,10}$ において、 $d_{m,10}=5$ のときは丸められうる値が2つ存在する。例えば $m=5$ のとき、0分に丸められる場合と10分に丸められる場合の2つが存在し、式(2)の性質上、これらの値へ丸められる確率は等しい。よって、この場合はそれぞれ丸められうる値への丸め誤差発生確率 $\tilde{p}_{m,10}$ は以下のようになることに注意されたい。

$$\tilde{p}_{m,10} = \frac{1}{2}p_{m,10} = \frac{1}{2}a_{10} \quad (6)$$

$p_{m,30}$ における $d_{m,30}=15$ のとき、 $p_{m,60}$ における $d_{m,60}=30$ のときも式(6)と同様の取り扱いを行う。

(4) 最尤法によるパラメータ推定の定式化

回答者が真に m 分に出発した確率を g_m をとおくと、 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。

$$q_m = g_m \cdot [1 - p_{m,5} - p_{m,10} - p_{m,15} - p_{m,30} - p_{m,60}] \quad (7)$$

式(7)の第2項はどの値にも丸められない確率を示している。例えば、 $m=8$ のとき、回答者が8分と回答する確率は、回答者が真に8分に出発した確率と、回答者がどの値にも丸めて回答しない確率の積で表される。

一方、 m が丸められうる値のとき、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。

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

式(7)の第1項および第2項は15分の周辺の時刻で真に出発し、かつ15分丸められて回答された確率の和を表

している。第1項は15分単位、第2項は5分単位でそれぞれ丸められる確率の和である。第3項は式(7)と同様に、真に15分に出発した確率と、どの値にも丸められて回答されない確率の積を示している。なお、 $m=15$ のとき $p_{m,5} = p_{m,15} = 0$ であるため、式(8)の第3項は式(7)の右辺と等価である。

式(7)および式(8)における、回答者が真に m 分に出発した確率 g_m の取り扱いについては、先行研究¹⁵⁾にならい「トリップの出発は一樣に分布している」という仮定を置く。つまり、 $g_m = 1/60$ である。

このとき、式(7)および式(8)を用いて、 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。

$$L = \prod_{m=0}^{59} q_m^{N_m} \quad (9)$$

対数尤度関数は以下の通りである。

$$\ln L = N_0 \ln q_0 + N_1 \ln q_1 + \dots + N_{59} \ln q_{59} \quad (10)$$

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。

$$0 \leq q_m, p_{m,5}, p_{m,10}, p_{m,15}, p_{m,30}, p_{m,60} \leq 1 \quad (11)$$

以上より、丸め誤差発生確率の推定は式(11)を満たしながら式(10)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。モデル推定にあたって、「丸め誤差は発生せず、真に出発した時刻と報告される時刻は同一である」という帰無仮説を置き、初期対数尤度を以下のように設定した。

$$\ln L_0 = N_0 \ln(1/60) + N_1 \ln(1/60) + \dots + N_{59} \ln(1/60) \quad (12)$$

(5) モデル推定結果

推定結果を表-2に示す。サンプルサイズが大きいことへの留意は必要であるが、すべてのパラメータで1%水準で有意となった。また、初期対数尤度(L_0)と対数尤度(L)を用いて尤度比検定を行うとすべての都市で1%水準で有意差が認められ、先述の帰無仮説は棄却されることが示された。多くの都市で支配的なのは a_{30} または a_{60} であり、30分または60分単位の丸め誤差発生確率が高いことがわかる。熊本PT調査では a_5 が支配的であった¹⁶⁾が、本研究の対象都市は5分単位の回答は少なく、対照的な結果となった。また、実際の出発時刻と丸められうる値との差による抵抗の意味を持つ b_5 については、どの都市も0に近く、丸め誤差発生に与える影響は殆ど見られない。一部の都市では、 b_5 が正の値として推定された。 b_5 の性質を考慮すると負の値であることが望ましいが、影響は限定的であると言える。

表-2 丸め誤差発生確率モデル推定結果

都市	トリップ数	最尤推定値					b_5	LL	L0	
		a_5	a_{10}	a_{15}	a_{30}	a_{60}				
1	Manila	471,035	0.079	0.063	0.119	0.359	0.380	-0.003	-619,805	-1,928,580
2	Chengdu	67,961	0.135	0.366	0.045	0.371	0.078	-0.006	-131,615	-278,256
3	Managua	54,137	0.073	0.116	0.135	0.345	0.318	-0.001	-82,277	-221,656
4	Belem	59,529	0.043	0.091	0.128	0.408	0.329	-0.003	-77,762	-243,732
5	Bucharest	143,311	0.098	0.129	0.158	0.368	0.245	-0.003	-229,418	-586,765
6	Cairo	268,360	0.014	0.033	0.131	0.297	0.508	0.006	-284,120	-1,098,758
7	Kuala Lumpur	218,460	0.107	0.119	0.208	0.361	0.203	-0.002	-361,286	-894,451
8	Ho Chi Minh	262,372	0.205	0.167	0.109	0.337	0.186	-0.006	-482,305	-1,074,241
9	Hanoi	188,949	0.108	0.107	0.175	0.418	0.192	-0.003	-299,978	-773,622
10	Nairobi(2005)*	39,167	0.012	0.064	0.112	0.412	0.400	(0.000)	-43,976	-160,363
11	Lima	299,133	0.134	0.157	0.135	0.328	0.241	-0.009	-476,248	-1,106,935
12	Dar es Salaam*	26,775	0.062	0.059	0.107	0.355	0.414	(0.000)	-33,017	-109,626
13	DaNang	50,509	0.143	0.153	0.207	0.320	0.181	-0.013	-92,087	-206,801
14	Ulaanbaatar	38,015	0.086	0.337	0.024	0.233	0.322	-0.007	-62,384	-155,647
15	Dhaka	153,654	0.038	0.072	0.062	0.378	0.452	-0.003	-167,430	-629,072
16	Nairobi(2013)	37,777	0.040	0.103	0.087	0.383	0.387	-0.008	-47,765	-154,672
17	Maputo	52,668	0.091	0.139	0.112	0.330	0.323	-0.003	-80,906	-215,641
18	Phnom Penh	86,884	0.050	0.118	0.046	0.418	0.369	-0.004	-107,773	-355,733
19	Karachi	217,875	0.048	0.026	0.087	0.359	0.476	0.002	-225,744	-892,055

(注) *の都市は報告されていない出発時刻の分が多く、パラメータが推定できないため、 b_5 を0に固定して推定した。

3. 主成分分析とクラスタ分析による回答傾向の類型化

前章では、本研究の対象とする19都市について丸め誤差発生確率モデルを用いてパラメータを推定し、丸め誤差の発生傾向について比較を行った。ただし、パラメータ推定値の大小の比較だけでは都市間の傾向の違いを俯瞰的に把握することは難しい。そこで本章では、主成分分析とそれを用いたクラスタ分析を行うことで対象19都市がどのように分類できるか、類型化を行うことを目的とする。

(1) 丸め誤差発生確率に関する主成分分析

前章で推定した、丸め誤差発生確率モデルのパラメータを用いて主成分分析を行う。本手法では、0分～59分の60個あるトリップ出発時刻の分布を6次元で考えられるため、より結果が解釈しやすくなることが期待できる。分析結果を表-3に示す。第1主成分では主に a_5 、 a_{10} 、 a_{60} が影響を与えていることから、60分単位の丸め誤差発生確率と5分または10分単位での比較的細かな丸め誤差発生確率との関係と考えられる。一方、第2主成分は a_{10} 、 a_{15} 、 a_{30} の影響が大きいことから、10分単位での丸め誤差発生確率と15分単位での丸め誤差発生確率との関係と考えられる。固有値が1以上となったのは第2主成分までで、累積寄与率が73.9%となった。

(2) クラスタ分析による類型化

前節で示した主成分分析をもとに、対象19都市に関するクラスタ分析を行う。クラスタ分析には大きく分けて階層クラスタリングと非階層クラスタリングの2種類が存在するが、本研究ではK-means法による非階層クラスタ分析を用いた。K-means法による非階層クラスタリングを行う場合、クラスタ数の決定方法でしばしば議論が行われるが、本研究ではクラスタリング結果を良好に説明できるクラスタ数である4クラスタでの分析とした。

図-3にクラスタリング結果を示す。横軸に主成分分析の第1主成分、縦軸に同じく第2主成分をとっている。

第1主成分においては、クラスタ1とクラスタ4で傾向が分かれる結果となった。つまり、60分単位の丸め誤差発生確率と5分または10分単位での比較的細かな丸め誤差発生確率との関係において、これらのクラスタ間で大きく傾向が変わることを示唆している。また、第2主成分においては、クラスタ2のUlaanbaatarを除き傾向に大きな傾向は見られない。クラスタ4では、クラスタ内でのばらつきがやや大きいことも見て取れる。

クラスタリング結果と地理的要因との関係では、(i)ベトナムのDa Nang, Hanoi, Hoh Chi Minhの3都市が同じクラスタ4に分類され、距離が近いKuala Lumpurも同クラスタに分類されている、(ii)中国とモンゴルの東アジア2国が同じクラスタ2に分類されている、(iii)クラスタ1では東南アジア～東アフリカの広い地域で、主にインド洋周辺の都市が分類されている、などの傾向が見られるが、図-4に示すように地理的要因と丸め誤差発生傾向の明確な関係は見られなかった。クラスタ3に分類された6都市や、中南米のManagua, BelemとLimaの関係、東ヨーロッパのBucharestとクラスタ4の他都市との関係など、地理的要因だけでなく文化的な背景や政治体制も含めた分析を今後の課題としたい。

(3) 各クラスタのトリップ出発分の特徴

ここでは、前節で行ったクラスタごとに実際のトリップ出発時刻の分布を整理し、その分布から丸め誤差発生に関する傾向を整理する。加えて、本研究で用いた丸め誤差発生確率による非階層クラスタリングが良好な分類結果を示しているか確認することも目的とする。

図-5はクラスタ1に分類された5都市のトリップ出発時刻の分布を示したものである。後述する他クラスタに比べ、00分と回答されたトリップ数が多く、これと30分を除いた他の丸められる値での回答が極端に少ない。差は僅かであるが、15分、45分での回答が5分単位や10分単位での回答に比べて多い傾向は見られる。

図-6ではクラスタ2に分類された2都市のトリップ出発時刻の分布を示す。(2)でも述べたように、この2都市は東アジアのChengdu(中国)、Ulaanbatar(モンゴル)であり、

表-3 丸め誤差発生確率の主成分分析結果

	PC1	PC2	PC3	PC4	PC5	PC6	
固有値	1.726	1.206	0.939	0.666	0.493	0.021	
累積寄与率	0.496	0.739	0.886	0.959	1.000	1.000	
因子負荷量	a_5	0.513	-0.176	0.133	-0.180	-0.759	0.282
	a_{10}	0.468	0.405	-0.140	-0.300	0.469	0.537
	a_{15}	0.048	-0.719	0.480	0.045	0.389	0.311
	a_{30}	-0.156	-0.472	-0.825	-0.084	-0.027	0.253
	a_{60}	-0.526	0.255	0.136	0.348	-0.219	0.686
	b_5	-0.463	0.007	0.181	-0.864	-0.070	0.023

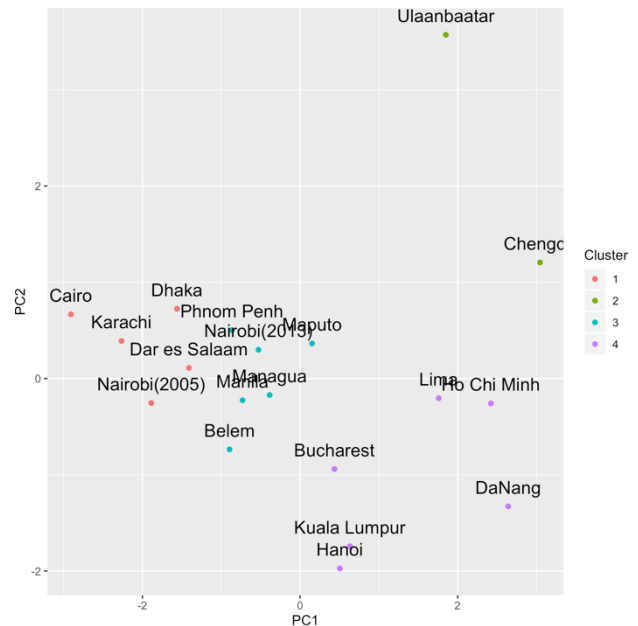


図-3 クラスタリング結果

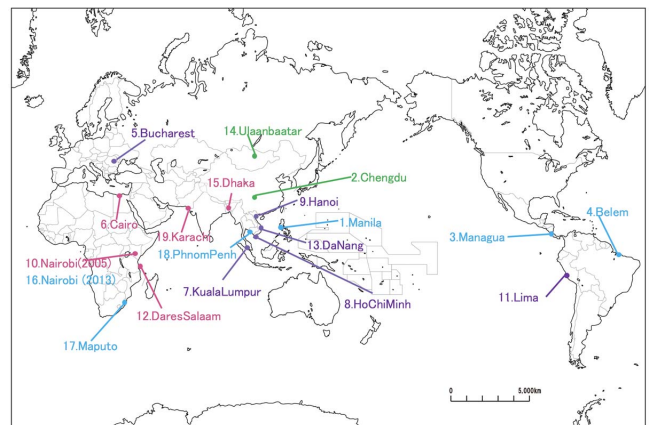
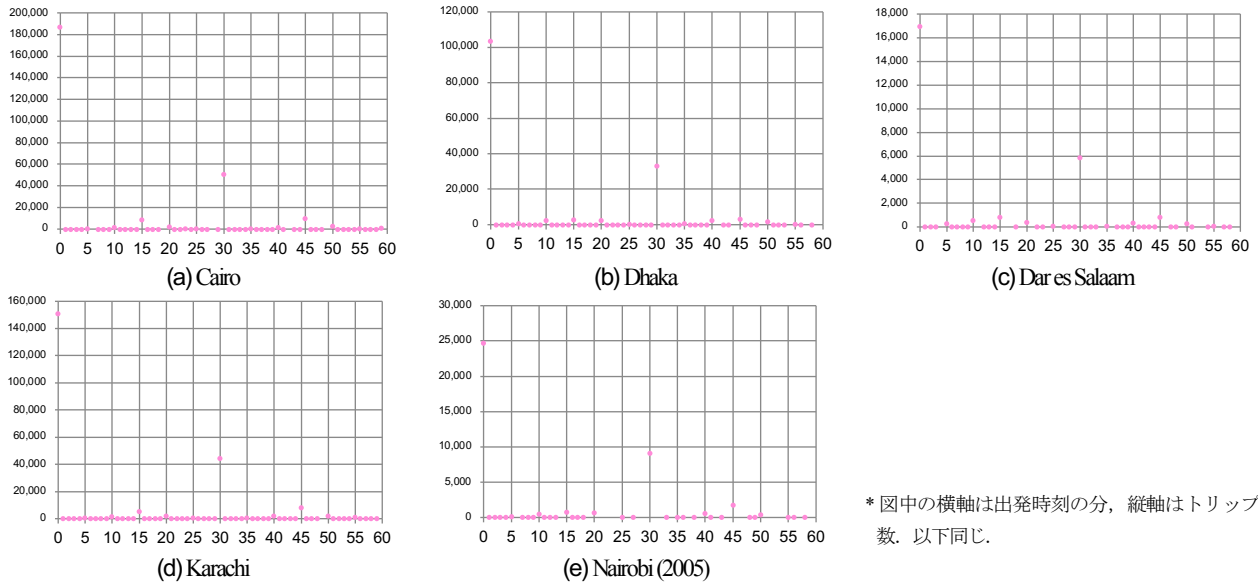


図-4 クラスタリング結果のプロット図

10分単位での回答が多いことが特徴的である。図-6の右端に、参考として熊本PT調査のトリップ出発時刻の分布を再掲する。2都市と距離の近い日本の調査でも同様の傾向が見られている。本研究では、調査主体が違うこと、データの傾向および丸め誤差発生確率モデル推定結



* 図中の横軸は出発時刻の分，縦軸はトリップ数。以下同じ。

図-5 トリップ出発分の分布(クラスタ 1)

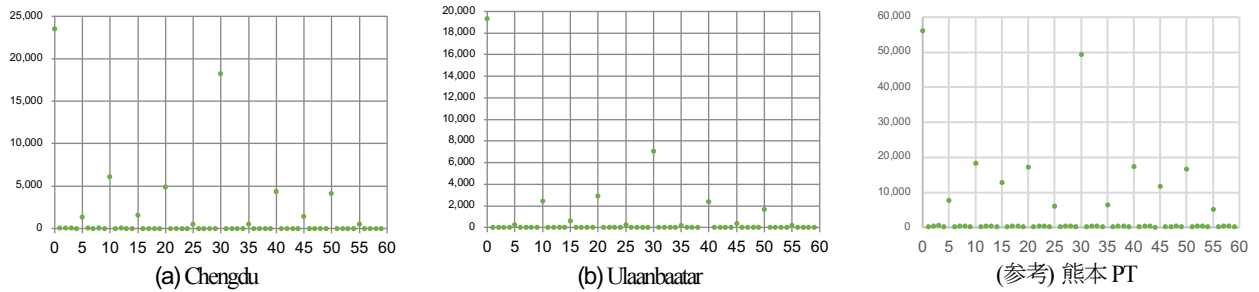


図-6 トリップ出発分の分布(クラスタ 2)

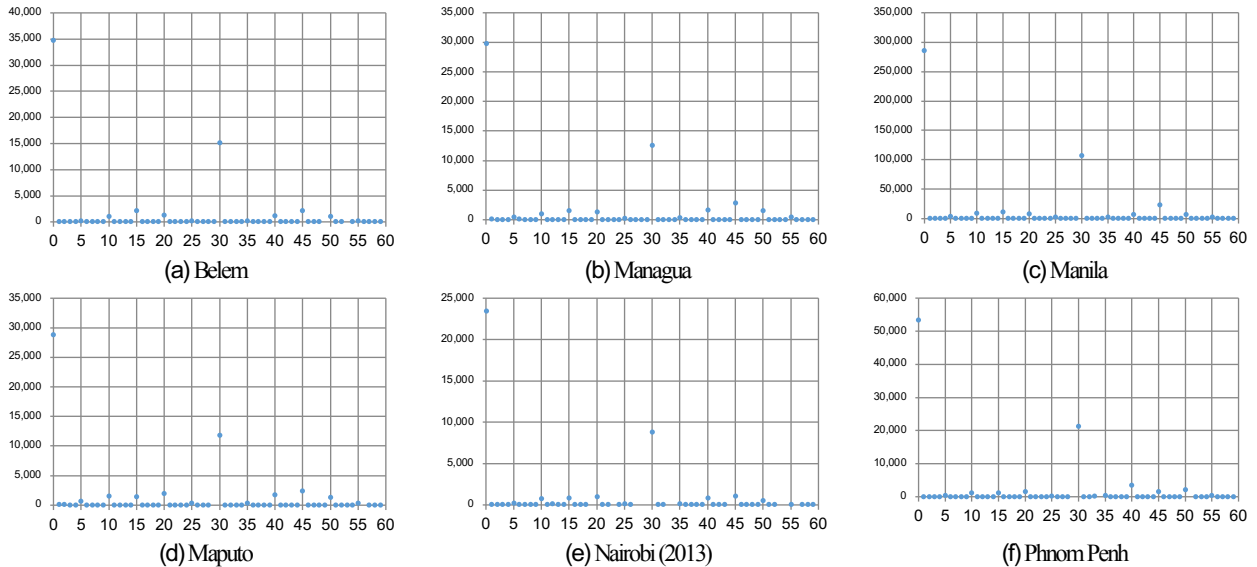


図-7 トリップ出発分の分布(クラスタ 3)

果に大きな差が見られていることから、熊本PTはクラスタリングの対象から外しているが、図-6の結果より熊本もこのクラスタに入ることが示唆されている。

図-7ではクラスタ3に分類された6都市のトリップ出発時刻の分布を示す。クラスタ1に似ており、00分と回答されたトリップが多い傾向にある。クラスタ1との相違

点は、5分～15分単位での回答が一定数見られることである。加えて、(a) Belem, (b) managua, (c) Manilaの3都市は5分、10分単位での回答に比べ15分単位での回答が僅かに多く、欧米諸国の調査によく見られる傾向¹³⁾を示していることも興味深い。

図-8ではクラスタ4に分類された5都市のトリップ出発

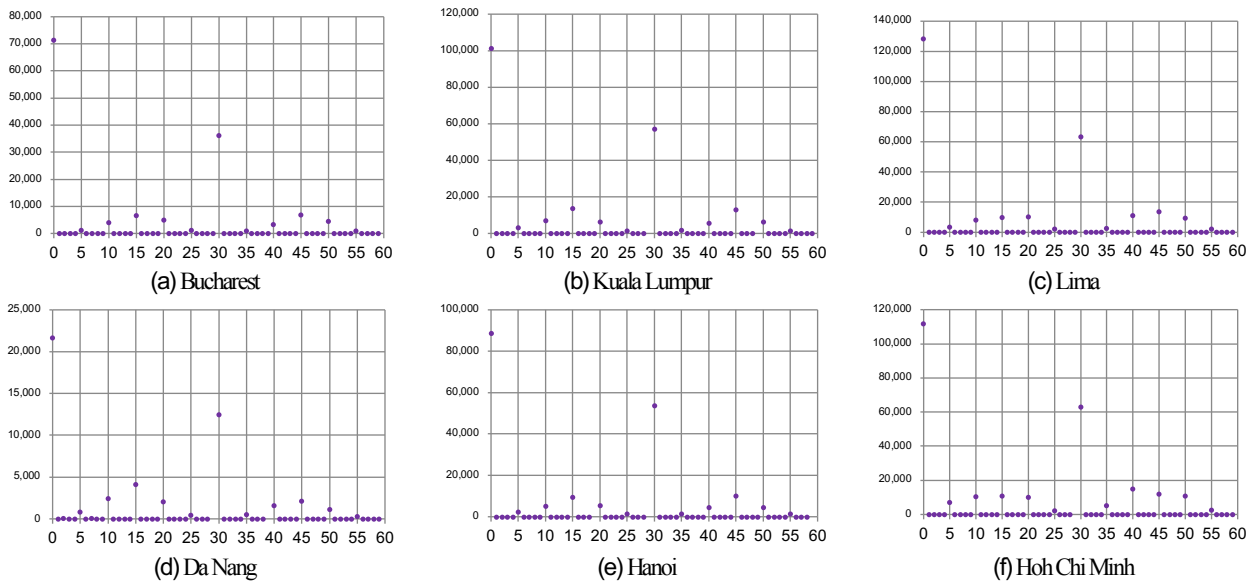


図-8 トリップ出発分の分布(クラスタ 4)

時刻の分布を示す。クラスタ3に比べ00分の単位での回答が少なく、比べて5分～15分単位での回答が多いことが特徴である。加えて、(f) Hoh Chi Minhでやや弱いですが、クラスタ3で述べた15分単位での回答が、クラスタ4でより顕著であることが見て取れる。

最後に、これらの特徴を表-5にまとめる。前章で行った主成分分析においても、第1主成分が「60分単位の丸め誤差発生確率と5分または10分単位での比較的細かな丸め誤差発生確率との関係」、第2主成分が「10分単位での丸め誤差発生確率と15分単位での丸め誤差発生確率との関係」と考えられ、実際のトリップ出発時刻の分布からもこれらの視点から良好にクラスタリングできていると考えられる。

また、今回構築したモデルおよびクラスタリングでは表現できていないが、クラスタ2の(e) Nairobi (2013)、(f) Phnom Penh、クラスタ4の(c) Lima、(f) Hoh Chi Minhの4都市は、30分へ近づくに従って5分～15分単位の回答が少しずつ増えていることは非常に興味深い。この傾向は、本研究のモデルでの推定が難しいため、傾向を適切に捉えられるモデルへの改良を今後の展望としたい。

表-5 各クラスタの特徴のまとめ

クラスタ	特徴
1	<ul style="list-style-type: none"> ・00分の回答が多く、00分、30分以外の回答は少ない ・5分～15分単位では、15分単位での回答が僅かに多い
2	<ul style="list-style-type: none"> ・15分単位に比べ10分単位での回答が多い
3	<ul style="list-style-type: none"> ・00分単位での回答が多いが、5分～15分単位の回答も一定数見られる ・15分単位での回答の傾向が僅かに見られる
4	<ul style="list-style-type: none"> ・00分の回答は比較的少なく、5分～15分単位での回答も多く見られる ・15分単位での回答の傾向が比較的高い

4. 旅行時間から見た各クラスタの特徴

前章までは、トリップ出発時刻の「分」に着目して丸め誤差の発生確率や傾向の類型化を行った。先行研究でも述べているが、トリップ到着時刻を研究の対象から外しているのは、例えば通勤トリップにおける始業時刻など、到着時刻の目安となる時間が存在する可能性があるためである。ここで、例えば「普段の通勤時間は20分」など、出発/到着時刻ではなく旅行時間はある程度正確に覚えている、という考え方もあろう。そこで本章では、「回答者は移動に要した時間はある程度正確に覚えており、かつ到着時刻の目安となる時間が存在するため、それらが出発時刻の丸め誤差発生に影響を及ぼす」との仮説を設定する。その仮説に基づき、旅行時間とトリップ出発時刻の回答傾向の関係について知見を整理する。

旅行時間に関しては図5～図8より、1分単位での回答はほぼ見られないことから、5分ごとの累積度数分布で比較することとした。本データ群では、旅行時間が0分のトリップが存在する都市と存在しない都市があるが、旅行時間に関する分析を目的とするため0分のトリップは除外した。加えて、超長時間のトリップが多く存在する都市もあるが、超長時間のトリップでは丸め誤差の発生確率が特に高くなると考えられるため、本研究では旅行時間が120分までのトリップを分析対象とした。都市によっては120分を超えるトリップを含めるべきである可能性もあるが、対象とする旅行時間の精査は今後の研究課題としたい。

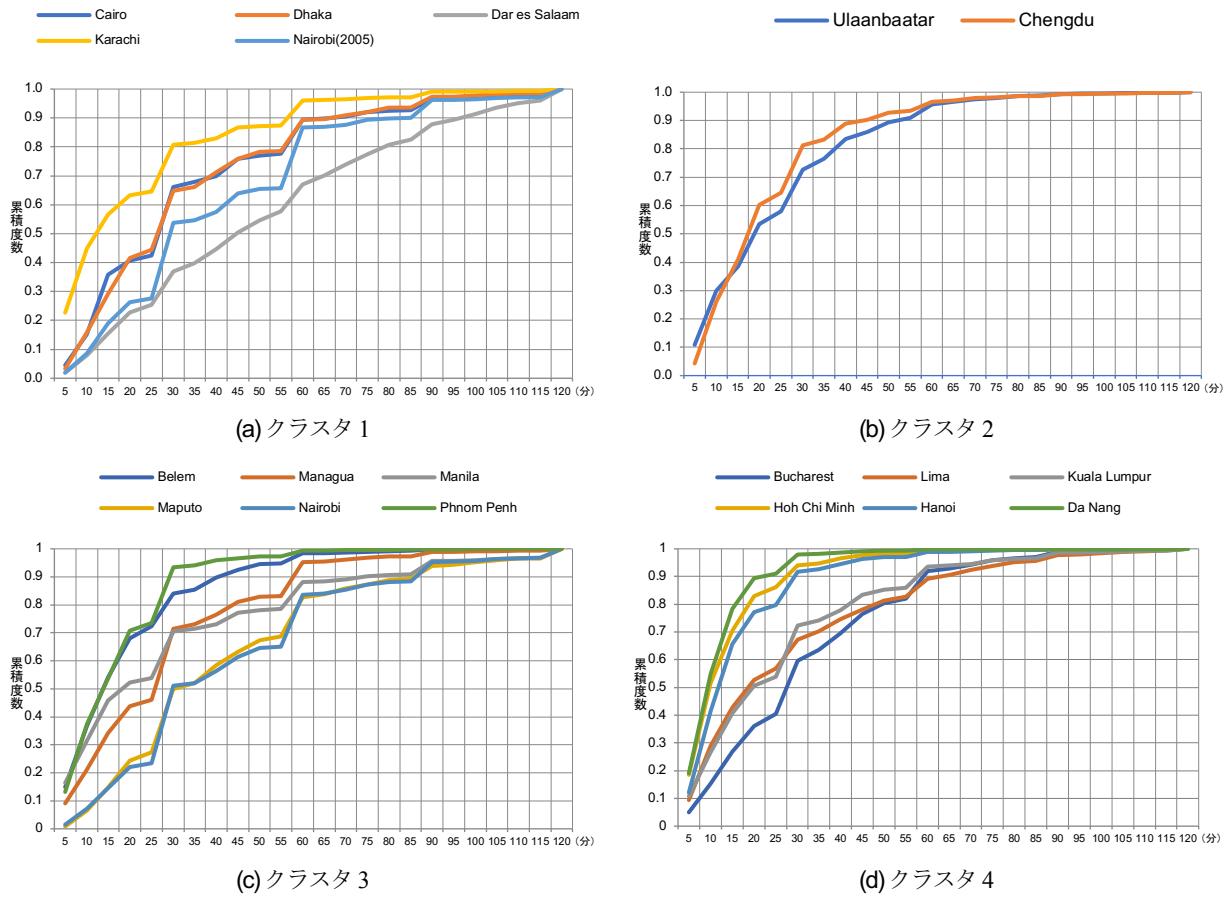


図-8 クラスタ別 旅行時間の累積度数分布

(1) 各クラスタにおける旅行時間の特徴

図-8にクラスタごとの旅行時間に関する累積度数分布を示す。ここでは、(i)全体の度数の傾きと(ii)1区間での度数の傾きに注目して考察を行う。クラスタ1は4クラスタ中もっとも全体の傾きが緩やかであり、長時間のトリップが多いクラスタと言える。1区間での度数の傾きでは、クラスタ1中の多くの都市で30分と60分の区間で傾きが急であり、90分でも同様の傾向が見られる。これらの都市では、移動が長時間の傾向かつ30分単位で考えられていることが結果から伺える。

クラスタ2では、階段状のような区間ごとの大きな傾きの変化は見られず、比較的なめらかな曲線を描いていることが特徴である。その中でも、10分、20分、30分の区間で傾きが急になっていることが見て取れる。また、全体の傾きも他と比べて急であり、これらのことからクラスタ2では比較的短時間かつ移動は10分単位で考えられていることが示唆される。

クラスタ3では、クラスタ1に比べて全体の傾きが急である都市が多いが、クラスタ1とよく似たパターンとなった。図-3のクラスタリング結果からも、クラスタ1とクラスタ3は近い関係にあることがわかる。ただし、クラスタ3ではBelem, Maputoを除く4都市で15分以内の1区

間ごとの傾きがクラスタ1の都市と比べてやや急である。クラスタ3の都市では、移動は30分単位で考えられているが、比較的短時間の移動も多いことが結果から伺える。

クラスタ4は、大きく2つのグループに分類できる。一方は20分以内の1区間の傾きをもっとも急である3都市で、同じベトナムのHoh Chi Minh, Hanoi, Da Nangとなったのは興味深い。これらの都市では短時間の移動が大きな割合を占めていることが分かる。もう一方は、クラスタ1, 3と同様の傾向を取るグループである。2つのグループに共通するのは、30分の区間での傾きが急であるが、60分の区間での傾きの変化がクラスタ1, 3と比較して緩やかであることが見て取れる。これらの都市では、移動が比較的短時間または30分程度と考えられていることが結果から推察される。

以上のように、トリップ出発時刻の回答傾向のタイプによって、旅行時間の分布にも様々な特徴が見られ、本章の冒頭で設定した「回答者は移動に要した時間はある程度正確に覚えており、かつ到着時刻の目安となる時間が存在するため、それらが出発時刻の丸め誤差発生に影響を及ぼす」という仮説を支持する結果も伺えた。ただし、クラスタ1とクラスタ3、クラスタ4のベトナム以外の3都市の違いなど、旅行時間だけでは類型をうまく説明でき

ない課題も残っている。トリップ手段別または目的別の分析を通して、本仮説の検証を今後の課題としたい。

(2) 実際の出発時刻の推定と都市間比較

最後に、本研究で用いた丸め誤差モデルから実際の出発時刻が推定できるため^{9,10}、実際の出発時刻の確率分布について都市間比較を行う。回答者が m 分と回答したとき、実際の出発時刻が n 分である確率を $P(n|m)$ とすると、ベイズの定理を用いて以下のように表される。

$$P(n|m) = \frac{P(m|n)P(n)}{P(m)} \quad (13)$$

右辺のうち、 $P(n)$ は実際の出発時刻が n 分である確率であり、これは3.での仮定に基づき $P(n) = 1/60$ である。 $P(m|n)$ は実際の出発時刻が n 分であるとき、回答者が m 分と回答する確率なので、これは丸め誤差発生確率に他ならない。 $P(m)$ は回答者が m 分と回答する確率なので、すでに2章で定式化されている。

表-6は、クラスタ1に分類されるCairoとクラスタ2に分類されるChengduの実際の出発時刻の累積確率を比較したものである。それぞれ00分に出発したと推定される確率は数%程度だが、実際の出発時刻が前後15分以内と推定される確率はChengduが89.3%に対し、Cairoが64.7%と大きく異なる。このような誤差が生まれる背景は、調査設計・実施手法よりも、回答の癖など回答者自身に起因するものが大きいと考えられるため、実際の出発時刻を正確に取得することは難しい。しかし、都市間比較を行う際は表-6のように確率分布が異なることが考慮できることが好ましく、この誤差を考慮したデータ比較手法の構築が望まれる。

表-6 実際の出発時刻の累積確率 (回答：00分)

誤差 実際の出発時刻		なし 00分*	10分以内 50分～10分	15分以内 45分～15分
確率	Cairo (クラスタ1)	2.4%	46.1%	64.7%
	Chengdu (クラスタ2)	4.8%	69.5%	89.3%

*厳密には、実際の出発時刻が00.0000分である確率は限りなく0に近いが、調査は1分単位での回答であるため、ここでは00分と記載する。

5. おわりに

本研究では、海外19都市の都市交通分野の開発調査における交通調査統合データを用いて、トリップ出発時刻の丸め誤差発生に関する傾向を分析した。結果を以下に示す。

- 1) 丸め誤差発生確率モデルを用いて対象都市の丸め誤差発生の傾向を分析し、60分または30分単位で丸め

て回答される確率が高いこと、熊本PT調査に比べ5分単位での回答は少ない傾向にあること、実際の出発時刻と丸められうる値との差が丸め誤差に与える影響は限定的であることを示した。

- 2) 1)で推定したパラメータを用いてクラスタリングし、対象都市を「(1)00分の回答が特に多い」「(2)15分単位に比べ10分単位の回答が多い」「(3)00分の回答が多いが、10～15分単位の回答も見られる」「(4)00分の回答が比較的少なく、10分単位に比べ15分単位の回答が多い」の4つに類型化できることを示した。
- 3) 上記類型と旅行時間の関係を分析し、(1)は長時間かつ30分単位での移動が多いこと、(2)は比較的短時間かつ10分単位での移動が多いこと、(3)は移動を30分単位で考えているが、比較的短時間の移動も見られること、(4)は短時間の移動または30分程度の移動が多く見られる、という知見を得た。
- 4) 異なるクラスタ間の2都市を例に実際のトリップ出発時刻の分布を推定し、00分と回答されたトリップで実際に00分に出発した確率に大きな差は見られないが、前後10分以内または前後15分以内の累積確率ではクラスタ間に大きな差が生じる可能性を示した。今後の課題として、4)における仮説を検証するにはトリップの性質を細かく分析する必要がある。具体的に、トリップの目的等によって到着時刻の目安となるものが存在しうるのかが変わってくると考えられるため、トリップ手段別、目的別または基礎属性別での分析を進めていきたい。加えて、今回は主にアジア、アフリカ、中南米の都市を対象としたが、欧米諸国の分析は十分でない。他データの取得によって対象都市を拡大することも今後の展望としたい。

謝辞：本研究の実施にあたりデータを提供頂きました独立行政法人国際協力機構に感謝致します。ただし、本稿に含みうる誤りの責は筆者のみにあります。本研究は、JSPS 科研費 JP18H01561 の支援を受けた成果の一部です。

参考文献

- 1) Tourangeau R.: Cognitive aspects of survey measurement and mismeasurement, *International Journal of Public Opinion Research*, Vol. 15, Issue 1, pp. 3-7, 2003.
- 2) Bricka S.G., Sen S., Paleti R., Bhat C.R.: An analysis of the factors influencing differences in survey-reported and GPS-recorded trips, *Transportation Research Part C*, Vol.21, pp.67-88, 2012.
- 3) Montini, L., S. Prost, J. Schrammel, N. Rieser-Schüssler, Axhausen K. W.: Comparison of Travel Diaries Generated from Smartphone Data and Dedicated GPS Devices, *Transportation Research Procedia*, Vol. 11, pp. 227-241, 2015.
- 4) Yamamoto, T., Madre, J.L., de Lapparent, M., Collet, R.: A random heaping model of annual vehicle kilometres

- travelled considering heterogeneous approximation in reporting, *Transportation*, pp.1-19, 2018.
- 5) Levinson D. M., Kumar A : A Multi-Modal Trip Distribution Model: Structure and Application, *Transportation Research Record*, Vol.1466, pp.124-131, 1994.
 - 6) Murakami E., Wagner D. P. : Can Using Global Positioning System (GPS) Improve Trip Reporting?, *Transportation Research Part C*, Volume 7, pp.149-165, 1999.
 - 7) Varotto S. F., Glerum A., Stathopoulos A., Bierlaire M., Longo G. : Mitigating the Impact of Errors in Travel Time Reporting on Mode Choice Modelling, *Journal of Transport Geography*, Volume 62, pp. 236-246, 2017.06.
 - 8) 上原一輝, 佐藤嘉洋, 円山琢也 : 交通調査の出発・到着時刻の丸め誤差と代理回答バイアス, 第 59 回土木計画学研究発表会講演集, 2019.06.
 - 9) Rietveld P : Rounding of Arrival and Departure Times in Travel Surveys: An Interpretation in Terms of Scheduled Activities, *Journal of Transportation Statistics*, Vol.5, Issue 1, pp. 71-82, 2001.
 - 10) 佐藤嘉洋, 円山琢也 : 交通調査の出発時刻の丸め誤差に関するモデル分析, 第 59 回土木計画学研究発表会講演集, 2019.06.
 - 11) Hyodo, T., Montalbo, C.M., Fujiwara, A. and Soehodho, S.; Urban travel behavior characteristics of 13 cities based on household interview survey data, *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, pp. 23-38, 2005.
 - 12) 熊本都市圏総合交通計画協議会 : 熊本都市圏都市交通マスタープラン, 2016.
 - 13) Kitamura, R: Time-of-day characteristics of travel: an analysis of 1990 NPTS data. In: Special Reports on Trip and Vehicle Attributes, 1990 NPTS Report Series, Publication No. FHWA-PL-95-033, US Department of Transportation, Federal Highway Administration, Washington, DC, February, pp 4-1-4-56, 1995.

(2019.10.4 受付)

INTERCITY COMPARISON OF ROUNDING OF DEPARTURE TIMES IN TRAVEL SURVEYS

Yoshihiro SATO and Takuya MARUYAMA