

PT 調査における所要時間の回答傾向の分析と補正手法の提案

愛媛大学 賛助会員 ○松本優里 愛媛大学大学院 正会員 倉内慎也 白柳洋俊

1. はじめに

パーソントリップ調査（PT 調査）は、昭和 42 年に広島都市圏で大規模に実施されて以来、これまでに全国各地で行われ、都市交通マスタープランの作成に活用されてきた。主な調査項目は、トリップの目的、OD、交通手段などであるが、これまで所要時間のデータはあまり活用されてこなかった。その理由として、PT 調査が粗いゾーンで区切られていることがあげられる。通常、ゾーン内人口の重心（以下、セントロイド）を起終点とするトリップとして扱うため、所要時間を精確に回答していたとしても、実際の距離とずれが生じてくる。また、PT 調査では、所要時間の真値と回答値にずれが生じることが想定され、特に 30 分単位、10 分単位などの丸め誤差が存在すると考えられる。そこで本研究は、PT 調査の所要時間の回答の報告誤差の特徴を分析し、補正することを目的とする。

2. 基礎集計

本研究で用いている PT 調査は 2007 年に実施された松山市のデータである。

まず、松山市の PT 調査と PP 調査の所要時間のグラフを図 1 に示す。PT 調査では 1 分単位の回答割合が低く、5 分、10 分、30 分などに集中していて、これが丸め誤差と考えられる。また、移動手段別に集計したものを図 2 に示す。公共交通機関を使っているトリップは、その他の移動手段のトリップよりも 1 分単位の回答が多い。交通手段以外にも、目的や個人属性によって回答傾向に違いが見られた。これらを説明要因としてモデルの作成を行う。

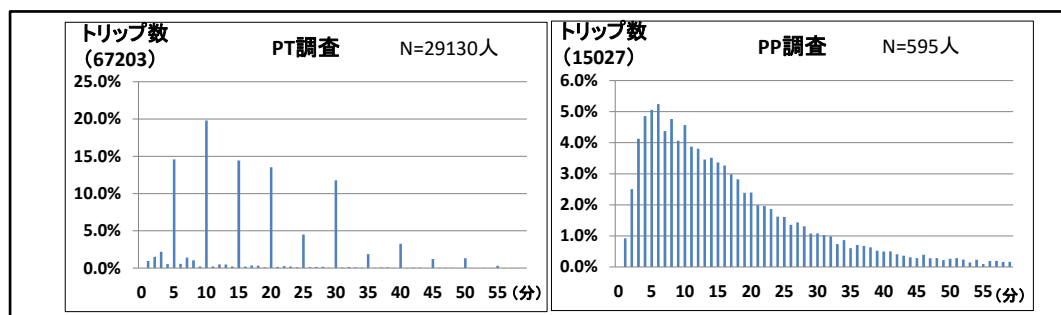


図 1 所要時間の回答傾向（PT 調査と PP 調査）

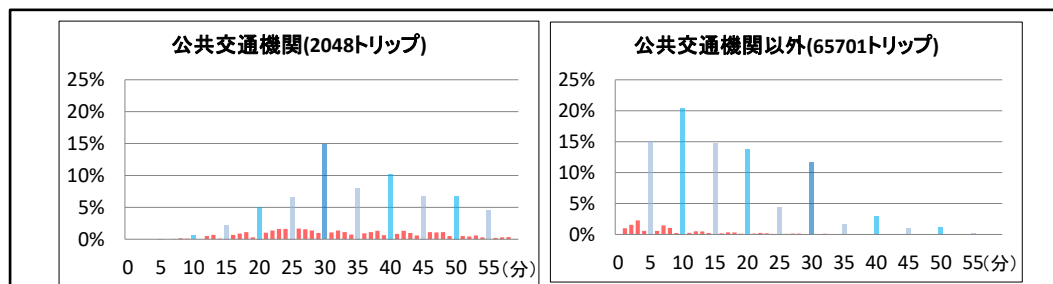


図 2 所要時間の回答傾向（移動手段別）

3. 報告誤差を考慮した所要時間モデル

丸め誤差の大きさが明らかな場合は、オーダードプロビットモデルの適用が考えられる。例えば、5 分単位に丸められると仮定した場合には、以下の式が得られる。ただし、 y_i' は丸め誤差を含む時刻の回答値、 y_i は真値として扱う。

$$\begin{aligned}
 y'_i &= 0 \text{ if } 0 \leq y_i < 2.5 \\
 &= 5 \text{ if } 2.5 \leq y_i < 7.5 \\
 &= 10 \text{ if } 7.5 \leq y_i < 12.5 \\
 &= 15 \text{ if } 12.5 \leq y_i < 17.5
 \end{aligned} \quad (1)$$

$$\begin{aligned}
 Pr(y'_i|z) &= \int_{u_{i-1}-\beta x_i}^{u_i-\beta x_i} \frac{1}{\sigma_\varepsilon} \phi\left(\frac{\varepsilon_i}{\sigma_\varepsilon}\right) d\varepsilon_i \\
 &= \Phi\left(\frac{u_i - \beta x_i}{\sigma_\varepsilon}\right) - \Phi\left(\frac{u_{i-1} - \beta x_i}{\sigma_\varepsilon}\right)
 \end{aligned} \quad (2)$$

ただし、 y'_i は丸め誤差を含む所要時間の回答値、 ϕ, Φ は標準正規確率密度関数、累積分布関数、 σ_ε は誤差項の分散、 u_i, u_{i-1} はそれぞれケース*i*の丸め誤差の上限値および下限値。

ここで、PT 調査データにはさまざまな大きさの丸め誤差が含まれており、全ての被検者が同一の丸め誤差に従うとは考えられず、ランダムな丸め誤差を考慮する必要がある。個々の回答値がいずれの丸め誤差に属するかを以下の式によって表す。

$$z_i^* = \alpha y_i + \Gamma x_i + \zeta_i \quad (3)$$

$$\begin{aligned}
 z_i &= 1 \text{ if } z_i^* < \theta_1, \\
 &= 2 \text{ if } 0 \leq z_i^* < \theta_2, \\
 &= 3 \text{ if } \theta_1 \leq z_i^* < \theta_3, \\
 &= 4 \text{ if } \theta_3 \leq z_i^*
 \end{aligned} \quad (4)$$

ただし、 α, Γ は未知パラメータおよび未知パラメータベクトル、 ζ_i は正規分布に従う誤差項、 $\theta_1, \theta_2, \theta_3$ は未知の閾値、 $z_i=1, 2, 3, 4$ 、はそれぞれ丸め誤差が1分単位、5分単位、10分単位、30分単位のセグメントに属することを示す。式(3)には丸め誤差を受ける前の所要時間である y_i が含まれており、所要時間が長いほど丸め誤差が大きくなる可能性を考慮している。また、尤度関数は以下の式で表される。

$$L = \prod_i \sum_{z \in Z_i} Pr(y'_i|z) Pr(z) \quad (5)$$

表1 所要時間補正モデルの推定結果

		回帰モデル回答値 推定値	提案モデル回答値 推定値
LOSデータ (内々)	定数項	1.902	2.01
	男性ダミー	0.0545	0.0531
	15歳以下ダミー	0.229	0.237
	65歳以上ダミー	0.0491	0.0516
	徒歩ダミー	0.396	0.398
	自転車ダミー	0.479	0.488
	二輪車ダミー	0.450	0.455
	自動車ダミー	0.466	0.470
	バス・鉄道ダミー	0.260	0.316
	ゾーン間所要時間	0.687	0.691
LOSデータ (内々以外)	ゾーン間距離	0.602	0.601
	徒歩ダミー	0.474	0.474
	自転車ダミー	0.528	0.533
	自動車ダミー	0.241	0.300
		所要時間	0.817
		男性ダミー	0.0326
		15歳以下ダミー	-0.0728
		65歳以上ダミー	-0.0546
		徒歩ダミー	1.65
		自転車ダミー	1.69
		二輪車ダミー	1.65
		自動車ダミー	1.66
		閾値 θ_1	2.63
		閾値 θ_2	4.82
		閾値 θ_3	6.04
		サンプル数	30000
		自由度修正済み決定係数	0.401
		標準偏差	0.592
			0.581

所要時間補正モデルの推定結果を表1に示す。左は、丸め誤差がないものとしてPT調査の回答値を被説明変数として回帰したもので、右は丸め誤差の存在を考慮し補正したうえでの提案モデルである。提案モデルでは、推定値の下の部分からどの丸め誤差に所属しているかを把握することができる。丸め誤差を考慮した提案モデルのほうが標準偏差はわずかながら小さくなっている。しかし、推定値を一つ一つ見ていくと、この2つのモデルにそこまでの差は見られなかった。

	1.96<t
	1.65<t<1.96
	t<1.65

4. 結論

最後にPP調査の観測値と所要時間の予測値を比較してモデルの精度検証を行った。相関係数、RMSEともに提案モデルのほうがわずかながら良い値になったが、それでもほとんど相関は見られなかった。その理由として、PPデータに異常値が含まれていることと、GPSの精度があまり良くないことがあげられる。今後は、元データのクリーニングが必要となる。