

観測規模と精度が異なるPT/PPデータを同時に用いた活動場所選択モデル

伊藤 創太¹・羽藤 英二²

¹学生非会員 東京大学大学院 工学系研究科都市工学専攻 (〒113-8656 東京都文京区本郷7-3-1)
E-mail: itoh@bin.t.u-tokyo.ac.jp

²正会員 工博 東京大学大学院 工学系研究科都市工学専攻 (〒113-8656 東京都文京区本郷7-3-1)
E-mail: hato@bin.t.u-tokyo.ac.jp

都心部での滞在活動の発生やその特性の分析は、交通管制や空間設計において不可欠であり、より高解像度かつ大容量のデータが取れるようになってきている中で、それらのデータを融合して分析することが可能となりつつある。本研究では、滞在活動の発生とその滞在時間について、パーソントリップ (PT) 調査とプローブパーソン (PP) 調査の観測精度の違いに注目しながら、滞在選択モデルをプロビットモデルで表現し、PTとPPの融合推定により分析を行う。分析から、PP・PTの観測の違いを見出し、空間別の需要発生に関する知見を得られる可能性を示した。

Key Words : combined estimation, probe person survey, person trip survey, data fusion

1. はじめに

プローブパーソン調査 (PP調査) は、被験者の行動データを取得可能な、GPS機能付端末とwebシステムを組み合わせた交通行動調査であり、個人の非集計の行動データをリアルタイムに、そして数日~数十日にわたる期間を通して取得することが可能である調査である。従前の観測手法よりも優れた解像度かつ、個人単位でのパネルデータとしての交通行動データの取得が可能となり、GPS機能付携帯電話やスマートフォンの技術進歩と普及が進む中で、発展が期待される。PP調査では、従前の調査では得ることが難しかった、トリップ経路の移動軌跡や、day-to-dayでの交通行動の変化が観測可能である。PP調査は調査における被験者の負荷を軽減することにもつながる。羽藤ら¹⁾は、PP調査で得られた加速度データを基に、隠れマルコフモデルやサポートベクターマシンを用いて交通機関や行動文脈を推定し、高い精度で識別可能であることを明らかにしている。

我が国での交通計画の策定においてこれまで基本とされていたのがパーソントリップ調査 (PT調査) である。PT調査は大サンプルでの調査であることから集計データとしての母集団代表性は高く、各都市圏で10年ごとに継続的にとられてきた貴重な調査データであることは間違いないが、調査については、トリップ・滞在の抜け落

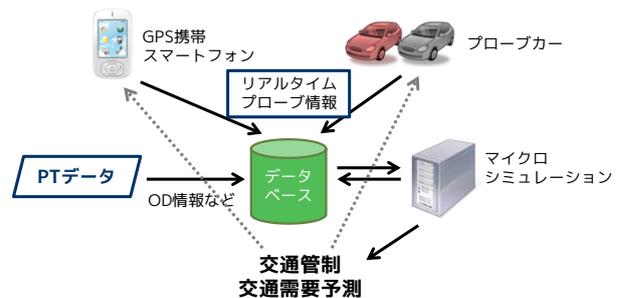


図-1 PP・PTデータの融合による交通管制

ちの発生や、ゾーン単位の集計による空間的集計誤差、また平日1日の横断面のデータのみにとどまっている、といった様々な問題点がかつてより指摘されている。

PP調査データも含め、多くのデータがとれるようになってきている中で、既存のPTデータなどのデータと、高精度で大量に取得できるようになりつつあるPPデータを融合させることにより、データベースとマイクロシミュレーションによる交通管制の実現の可能性が考えられる (図-1)。人の交通行動を予測するマイクロシミュレーションによって、OD交通量やマクロな交通配分をPTデータを基に行い、経路選択や需要の変動、従前の手法では詳しく分析することができなかったゾーン内トリップなどの微細な交通行動の分析はPPデータを用いて捉

え、分析を行うことなどができよう。

こうした背景を踏まえ、本研究では、PPデータとPTデータのデータ融合により、都心部での滞在活動の頻度や滞在時間といった特性を分析することを考える。このとき問題となるのがPT調査とPP調査における観測手法の違いによる観測精度の違いであり、その大きな要因として挙げられるのがPT調査における情報の抜け落ちの存在である。

本研究では、共通観測量を変数として、PT・PPデータを用いてゾーンごとの滞在選択モデルの融合推定を行い、滞在活動を推定する手法を試みる。このとき、PTデータでは滞在時間の長さが一定時間以下の滞在活動は調査時に欠損が発生していると仮定して、その閾値を境に短時間・長時間の2つの滞在選択モデルを構築する。PT・PPのデータ融合による滞在の推定により、PTデータだけでは考慮できなかった滞在観測データの抜け落ちを反映させる。

2. PT・PPデータの特徴

PT・PPの両調査では、滞在活動の場所・時間・個人属性については、ほぼ同様の項目について観測が行われる。ただし、その解像度や信頼性については差異があるといえる。例えば空間の集計単位をとっても、PT調査では滞在場所がゾーン単位で表されているのに対して、PP調査では位置座標での記録となっているように、仕様の違いがみられる。また、時間においても、PT調査では被験者の記憶に基づいて回答されていることから、活動の開始・終了時刻などは曖昧であったり、丸められたりして記録されている可能性があることに留意しなければならない。

また、PTデータとPPデータでは調査期間・対象の違いもある。PT調査では、平日1日の交通行動について対象とする母集団に関して横断的にデータが取得されているのに対して、PP調査では同一個人について数日～数十日にわたって複数日のデータが継続的に取られている。PT調査では各データはそれぞれ異なる個人の1日の交通行動である。それに対して、PP調査はPT調査に比べると被験者数は少数であるものの、各個人の調査期間分の毎日の行動が観測されているため、活動の頻度やday-to-dayの変動特性が分析可能である。

PTデータにおいて、最も留意しなければならない問題が、観測されたトリップ・滞在活動に抜け落ちが発生していることである。PT調査は紙媒体でのアンケート調査によって被験者が回答を行うことから、記憶が曖昧であったり、重要でないと考えられてしまう行動は、回答時に省かれている可能性がある。PP・PT調査の調査

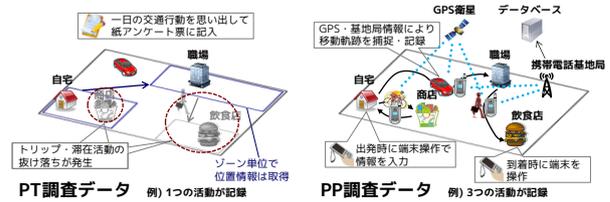


図-2 PT・PP調査の交通行動記録手法の概要

手法の違いを図-2に示す。PT調査では交通行動について、後から被験者の記憶を頼りに記録されるのに対し、GPSとweb技術を用いたPP調査ではリアルタイムに被験者の重い負担を強くない簡単な端末操作で記録される。既存の紙ベースでの調査票を用いた研究によると、井坪ら(2005)²⁾の研究では、GPS機能付携帯電話とwebダイアリーを用いた調査では、自動車のトリップ原単位が紙ベースの調査と比べて15%増加していることが確認されている。また、名取ら(2000)³⁾も、同一被験者によるPT調査とアクティビティダイアリー調査の回答を比較することで、PT調査でトリップの抜け落ちがあると考えられる被験者が30%に上ることが示されている。海外の事例においても、Stopher and Greaves(2010)⁴⁾はGPSによる調査と、紙ベースでの日誌調査を比較し、紙ベース調査では少なくとも20%のトリップの抜け落ちがあることを指摘している。

3. PT・PPデータ融合滞在選択モデル

(1) データ融合によるモデル

複数のデータを用いてモデル推定を行う例として、RP(Revealed Preference)データとSP(Stated Preference)データの同時推定が想起される。森川・山田(1993)⁵⁾は、RPデータとSPデータを融合したMNLモデルの推定において、誤差項のスケールを統一するスケールパラメータを導入する手法を提案している。RP・SPデータの融合推定の手法は、共通の変数が取得できている点、異なるサンプルデータを融合する点などからPP・PTデータの融合推定と共通する部分も多く、参考としながらPP・PTデータにおいても融合推定を試みる。

(2) PTデータの欠損を考慮した滞在场所選択モデル

本研究ではゾーンごとの滞在・活動の発生について、どの場所を訪れるかという訪問有無の選択を、プロビット選択モデルによって表す。このとき、モデルの推定に使用するデータとしてPTデータとPPデータを融合して使うことにより、データの信頼性の高いPPデータと、サンプル数の多いPTデータの双方の利点を生かした推定を行いたい。

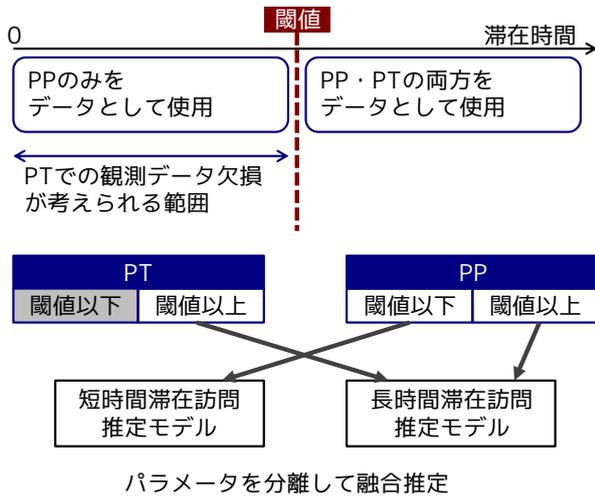


図-3 PP・PTデータを用いた融合推定の概要

ここで本研究では、PT調査では一定時間以下の滞在活動の観測において欠損が発生していると仮定する。このときの欠損が考えられる滞在時間の範囲の上限となる時間を閾値と呼ぶこととする（図-3）。PT調査において、閾値を下回る短時間滞在活動についての観測は欠損が生じるが、閾値以上の長時間の滞在活動の観測は正しく行われているとする。また、PP調査においては、逐次的に携帯端末を操作することにより記録が行われているため、いずれの滞在時間の場合も正しく記録が行われていると仮定する。

ゾーンごと、個人ごとの滞在の有無を表すモデルとして、プロビット選択型のモデルを用いる。プロビット選択モデルでは説明変数から潜在変数が定義され、潜在変数がある閾値を超える時にその値が観測される構造となっている。一般性を失わないため、ここでは0を超える時にそのデータは観測されるものとする。このとき、PTデータ・PPデータを用いた潜在変数は、個人*i*・目的地*n*について以下のように表す。

$$y_{in} = \beta x_{in} + \varepsilon_{in} \quad (3a)$$

y_{in} は滞在の発生の起こりやすさを表す潜在変数、 β はパラメータ、 x_{in} は説明変数である。 ε_{in} は目的地・個人ごとに独立な誤差項で、平均0、標準偏差 σ の正規分布に従う。ここでは一般性を失わないため、モデルの誤差項の標準偏差については、 $\sigma=1$ として基準化する。

(3) 短時間・長時間滞在の分離

PT・PPデータを用いた推定に関して、PTデータでは短時間の活動が抜け落ちている可能性があり、サンプリングバイアスが発生してしまうことを考慮し、閾値*T*を導入して短時間・長時間の滞在でそれぞれ別々にモデルのパラメータを設定する。

$$y_{in}^{short} = \beta^{short} x_{in} + \varepsilon_{in} \quad (3b)$$

$$y_{in}^{long} = \beta^{long} x_{in} + \varepsilon_{in} \quad (3c)$$

ただし、 y_{in}^{short} は、個人*i*がゾーン*n*に滞在した滞在時間 t_{in} が*T*以下の滞在活動の発生の起こりやすさ、 y_{in}^{long} は閾値*T*を超える滞在時間の滞在活動の起こりやすさを表し、それぞれのモデルのパラメータ β^{short} と β^{long} も別々にとることとする。

(4) 尤度関数の設定

尤度関数は、以下のように、短時間滞在についてのPPデータを用いたモデルの対数尤度と、PP・PTデータを用いた長時間滞在の対数尤度の和をとり、融合推定モデルの尤度関数とする。

短時間モデルの対数尤度を LL^{short} 、長時間モデルの対数尤度を LL^{long} とすると、モデル全体の対数尤度 LL は、

$$LL^{short} = \sum_{0 < t_{in} \leq T} \ln \left\{ 1 - \Phi \left(\frac{\beta^{short} x_{in}}{\sigma} \right) \right\} + \sum_{t_{in} > T, t_{in} = 0} \ln \left\{ \Phi \left(\frac{\beta^{short} x_{in}}{\sigma} \right) \right\} \quad (3d)$$

$$LL^{long} = \sum_{t_{in} > T} \ln \left\{ 1 - \Phi \left(\frac{\beta^{long} x_{in}}{\sigma} \right) \right\} + \sum_{t_{in} \leq T} \ln \left\{ \Phi \left(\frac{\beta^{long} x_{in}}{\sigma} \right) \right\} \quad (3e)$$

$$LL = LL^{short} + LL^{long} \quad (3f)$$

モデルのパラメータは、式(3f)の尤度を最大化する最尤推定法により計算する。

4. データの概要

ここでは、本研究での実証的分析で用いたデータの概要について説明する。

(1) PTデータ

本研究ではPTデータとして、2008年に東京都市圏で行われた第5回東京都市圏パーソントリップ調査を使用する。調査は東京都市圏の約3500万人を対象とし、抽出率は2.7%である。同調査では1,906,032件のマスターデータが得られており、居住都県・地域によって交通行動に差があることが考えられることから、横浜市内居住者を対象としたPPデータとの整合性を保つため、PTデータから横浜市内居住者データのみを抽出すると、253,737件が得られた。抽出したマスターデータより、滞在データを作成し、滞在场所については調査がゾーン単位での集計となっていることから、ゾーンのセントロイドで代表させ、計算を行った。横浜都心部の小ゾーンについては、図4

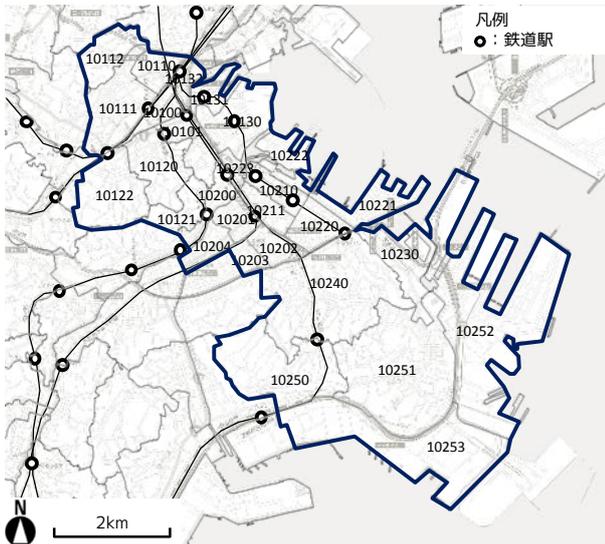


図4 横浜都心部のPT小ゾーン

に示す。

(2) PPデータ

PPデータは、2010年に横浜市で行われたプローブパーソン調査データを用いる。調査概要を表-1に示す。調査では、トリップの開始時・交通手段変更時・終了時に目的地での滞在目的や移動交通手段をGPS機能付携帯電話操作によって入力を行っている。GPS機能により出発地から到着地までの移動軌跡も取得される。PPデータでは、行動時にトリップの記録が行うが、トリップ・滞在の記録に誤りや修正がある場合であってもwebより被験者本人による修正が可能である。

なお、PTデータが平日1日の交通行動を尋ねる調査データとなっていることから、PPデータについても土曜・日曜・祝日を除いたデータのみを抽出して分析を行った。

PP調査により得られた個人別・ゾーン別の調査期間35日中の滞在回数を図-5に示す。図-5は各列が横浜都心部のPT小ゾーンを、各行がPP調査モニターを表し、PP調査モニターごとに、訪問場所をPT小ゾーンごとに集計し、調査期間中の滞在回数を集計した。表中の数字は各ゾーンで通勤・通学・業務目的を除いた非義務目的での滞在が観測された回数を示している。表を見ると、いずれも通勤や日常生活で横浜都心部を利用するモニターでありながら、非義務目的では全く利用しない人から、特定の場所に集中的に多く行く人、多くの場所にまんべんなく行く人と、個人差があることがわかる。また、ゾーンごとにみても、あまり滞在する人がいないゾーンや、多くの人が滞在し利用しているゾーンがあり、空間による滞在の特性の差異があることが考えられる。

表-1 横浜PP調査2010の概要

調査期間	2010年7月5日～8月8日 (35日間)
調査方法	携帯電話端末による行動記録調査とwebダイアリ調査
被験者数	40名
取得トリップ数	3617件
取得ロケーションデータ数	789,074件

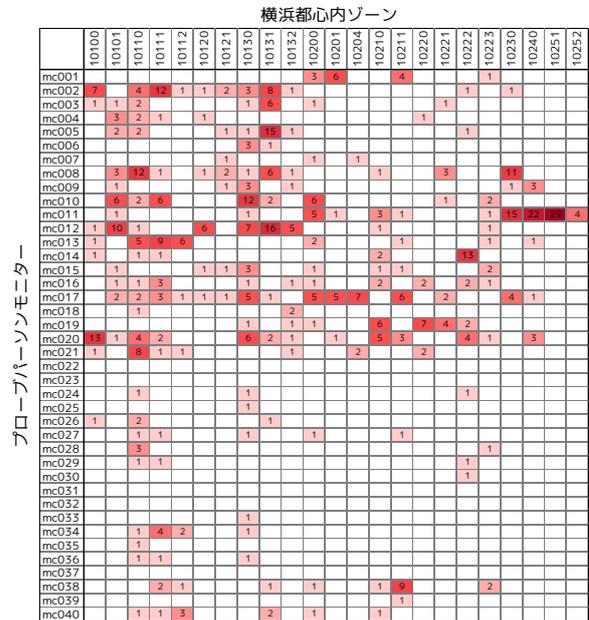


図-5 個人別・ゾーン別の調査期間中滞在回数

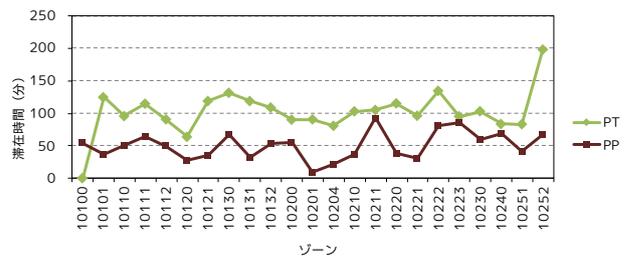


図-6 ゾーン別滞在時間集計

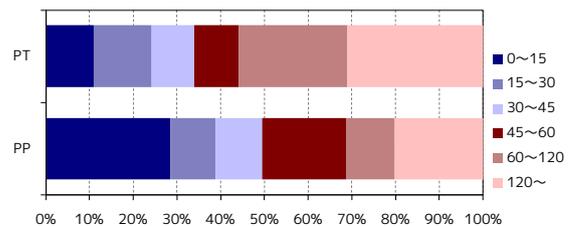


図-7 PT・PPの滞在時間構成割合比較

表-2 PP・PT融合モデル推定結果とPP・PTによる推定結果の比較

	融合推定 (閾値：30分)		PP (閾値：30分)		PT (閾値：30分)	
	推定値	t値	推定値	t値	推定値	t値
短時間滞在パラメータ						
定数項	-2.060	-15.74 **	-2.062	-15.75 **	-3.801	-102.99 **
男性	0.187	1.86	0.187	1.85	-0.044	-2.34 *
30歳代	0.198	1.56	0.198	1.56	0.217	4.97 :
40歳代	0.427	3.19 **	0.427	3.19 **	0.321	8.10 **
50歳代	0.402	2.38 *	0.402	2.38 *	0.300	7.24 **
60歳代以上	0.291	1.52	0.291	1.52	0.231	6.24 **
単身世帯	-0.209	-2.08 *	-0.209	-2.08 *	0.081	3.26 **
自動車保有	-0.305	-3.37 **	-0.304	-3.36 **	0.214	10.76 **
自宅からの距離(100m)	-18.627	-7.75 **	-18.579	-7.74 **	-0.673	-1.34
職場からの距離(100m)	-2.523	-4.56 **	-2.520	-4.55 **	-0.490	-4.29 **
ゾーン売場面積(ha)	0.023	3.73 **	0.023	3.73 **	0.047	34.33 **
長時間滞在パラメータ						
定数項	-3.239	-196.36 **	-2.053	-19.67 **	-3.261	-191.47 **
男性	-0.103	-10.01 **	0.139	1.47	-0.111	-10.59 **
30歳代	0.130	6.27 **	-0.154	-1.54	0.123	5.72 **
40歳代	0.095	4.84 **	-0.068	-0.62	0.099	4.93 **
50歳代	0.092	4.43 **	-0.081	-0.60	0.100	4.66 **
60歳代以上	0.175	10.57 **	0.281	1.81	0.185	10.82 **
単身世帯	0.100	7.12 **	-0.325	-3.33 **	0.093	6.39 **
自動車保有	0.119	10.65 **	-0.231	-2.96 **	0.126	11.05 **
自宅からの距離(100m)	-0.184	-1.07	-4.096	-5.51 **	-0.362	-1.72
職場からの距離(100m)	-0.840	-12.44 **	-3.712	-4.37 **	-0.767	-11.28 **
ゾーン売場面積(ha)	0.058	77.48 **	0.022	3.59 **	0.059	77.81 **
サンプル数	2354942		23000		2321942	
初期尤度 LL(0)	-1641332		-31884		-3218895	
最終尤度 LL(β)	-31789		-1629		-38604	
修正済み尤度比	0.981		0.949		0.988	

*:5%有意 ** :1%有意

(3) データの比較

PTデータとPPデータの融合推定の前に、2つのデータを比較を行う。図-6は、ゾーン別に滞在時間の平均をPT・PPの2つのデータで集計したものである。各ゾーンにより滞在時間の平均にはばらつきがみられ、PT・PPデータではほぼ同様の傾向を示している。しかしながら、その滞在時間平均値はほぼ全てのゾーンにおいてPPデータがPTデータを下回っている。この差の要因として、アンケート形式のPT調査では、短時間の滞在が抜け落ちていることが考えられる。

さらに、図-7に買い物・娯楽等の非義務目的滞在の滞在時間別構成割合を示す。図からPPデータでは15分以内の滞在が卓越しているのに対して、PTデータではそうした傾向はみられず、PTデータは短時間滞在において欠損が生じやすい可能性が考えられる。

5. PT・PP融合モデル推定結果

(1) 推定結果

推定結果を表-2に示す。表-2では閾値を30分で設定した結果を掲載し、比較としてPP・PTのデータをそれぞれ30分を閾値として短時間・長時間のモデルに用いた結

果を並べた。パラメータ推定値は正で値が大きくなるほど、その要因によって滞在の発生確率が高まることを意味する。融合推定の結果をみると、PP単独での推定よりも、データ数の多いPTデータを融合することで尤度比が向上していることがわかる。また、短時間部分における融合推定結果、PTの推定結果、PPの推定結果を比べると、定数項の値がPPの-2.062に対してPTでは-3.801と小さくなっていることから、短時間の滞在は全体としてPTでは過小に観測されていることが考えられる。

さらに個別の説明変数を見ると、滞在场所までの自宅からの距離と職場からの距離がPPではパラメータ値もPTより大きく、有意に効いていることから、PP調査では滞在场所までの距離の影響が強く表れていることがわかる。PTでは上記の2つの距離に関する説明変数のパラメータは卓越しておらず、目的地までの距離と活動需要との関係が十分に表すことができていないといえる。

(2) 閾値の設定

モデルでは、PTでの抜け落ちが発生する範囲を閾値によって設定したが、この閾値の取り方により推定の結果は変化すると考えられる。閾値によるパラメータや推定結果の尤度のふるまいを考察することで、観測の抜け

表-3 PP・PT融合推定モデル閾値による結果の変化

	融合推定 (閾値：15分)		融合推定 (閾値：30分)		融合推定 (閾値：45分)		融合推定 (閾値：60分)	
	推定値	t値	推定値	t値	推定値	t値	推定値	t値
短時間滞在パラメータ								
定数項	-2.296	-11.90 **	-2.060	-15.74 **	-1.952	-16.12 **	-2.155	-18.56 **
男性	0.218	1.68	0.187	1.86	0.180	1.89	0.171	1.93
30歳代	0.316	1.64	0.198	1.56	0.173	1.49	0.268	2.38 *
40歳代	0.636	3.20 **	0.427	3.19 **	0.412	3.33 **	0.486	4.13 **
50歳代	0.641	2.61 **	0.402	2.38 *	0.415	2.69 **	0.441	3.10 **
60歳代以上	0.646	2.47 *	0.291	1.52	0.255	1.42	0.432	2.59 *
単身世帯	-0.199	-1.53	-0.209	-2.08 *	-0.224	-2.35 *	-0.213	-2.37 *
自動車保有	-0.426	-3.64 **	-0.305	-3.37 **	-0.315	-3.70 **	-0.274	-3.62 **
自宅からの距離(100m)	-28.683	-6.72 **	-18.627	-7.75 **	-18.210	-8.29 **	-10.547	-8.32 **
職場からの距離(100m)	-2.343	-3.45 **	-2.523	-4.56 **	-2.826	-5.22 **	-2.806	-5.46 **
ゾーン売場面積(ha)	0.022	2.82 **	0.023	3.73 **	0.019	3.20 **	0.020	3.60 **
長時間滞在パラメータ								
定数項	-3.222	-203.13 **	-3.239	-196.36 **	-3.254	-190.97 **	-3.268	-184.62 **
男性	-0.100	-10.27 **	-0.103	-10.01 **	-0.100	-9.29 **	-0.099	-8.61 **
30歳代	0.158	8.03 **	0.130	6.27 **	0.110	5.06 **	0.095	4.18 **
40歳代	0.124	6.66 **	0.095	4.84 **	0.072	3.52 **	0.047	2.18 *
50歳代	0.125	6.31 **	0.092	4.43 **	0.065	2.99 **	0.041	1.76
60歳代以上	0.183	11.50 **	0.175	10.57 **	0.158	9.27 **	0.137	7.73 **
単身世帯	0.100	7.55 **	0.100	7.12 **	0.096	6.42 **	0.095	5.96 **
自動車保有	0.136	13.01 **	0.119	10.65 **	0.106	9.02 **	0.092	7.33 **
自宅からの距離(100m)	-0.141	-0.88	-0.184	-1.07	-0.336	-1.66	-0.358	-1.65
職場からの距離(100m)	-0.818	-12.99 **	-0.840	-12.44 **	-0.849	-11.89 **	-0.887	-11.57 **
ゾーン売場面積(ha)	0.059	82.75 **	0.058	77.48 **	0.058	72.86 **	0.056	67.10 **
サンプル数	2354942		-2354942		-2354942		-2354942	
初期尤度 LL(0)	-1641332		-1641332		-1641332		-1641332	
最終尤度 LL($\hat{\beta}$)	-35334		-31789		-28782		-25656	
修正済み尤度比	0.978		0.981		0.982		0.984	

*:5%有意 **:1%有意

表-4 PP・PT融合推定モデル混合率による推定結果の比較

	PP,PT 混合率 1:1 (閾値：30分)		PP,PT 混合率 1:10 (閾値：30分)		PP,PT 混合率 1:101 (閾値：30分)	
	推定値	t値	推定値	t値	推定値	t値
短時間滞在パラメータ						
定数項	-2.060	-15.74 **	-2.060	-15.74 **	-2.060	-15.74 **
男性	0.187	1.86	0.187	1.86	0.187	1.86
30歳代	0.198	1.56	0.198	1.56	0.198	1.56
40歳代	0.427	3.19 **	0.427	3.19 **	0.427	3.19 **
50歳代	0.402	2.38 *	0.402	2.38 *	0.402	2.38 *
60歳代以上	0.291	1.52	0.291	1.52	0.291	1.52
単身世帯	-0.209	-2.08 *	-0.209	-2.08 *	-0.209	-2.08 *
自動車保有	-0.305	-3.37 **	-0.305	-3.37 **	-0.305	-3.37 **
自宅からの距離(100m)	-18.627	-7.75 **	-18.627	-7.75 **	-18.627	-7.75 **
職場からの距離(100m)	-2.523	-4.56 **	-2.523	-4.56 **	-2.523	-4.56 **
ゾーン売場面積(ha)	0.023	3.73 **	0.023	3.73 **	0.023	3.73 **
長時間滞在パラメータ						
定数項	-2.640	-36.60 **	-3.148	-70.54 **	-3.239	-196.36 **
男性	0.056	1.07	-0.051	-1.72	-0.103	-10.01 **
30歳代	0.088	1.14	0.151	2.79 **	0.130	6.27 **
40歳代	-0.055	-0.63	0.079	1.48	0.095	4.84 **
50歳代	-0.033	-0.32	0.062	1.05	0.092	4.43 **
60歳代以上	0.039	0.48	0.161	3.57 **	0.175	10.57 **
単身世帯	0.012	0.20	0.103	2.72 **	0.100	7.12 **
自動車保有	0.026	0.48	0.053	1.70	0.119	10.65 **
自宅からの距離(100m)	-1.841	-3.50 **	0.284	0.97	-0.184	-1.07
職場からの距離(100m)	-2.861	-5.28 **	-1.388	-6.46 **	-0.840	-12.44 **
ゾーン売場面積(ha)	0.032	6.97 **	0.054	24.49 **	0.058	77.48 **
サンプル数(PP)	23000		23000		23000	
サンプル数(PT)	23000		230000		2321942	
初期尤度 LL(0)	-47827		-191309		-1641332	
最終尤度 LL($\hat{\beta}$)	-2034		-4530		-31789	
修正済み尤度比	0.957		0.976		0.981	

*:5%有意 **:1%有意

落ちのメカニズムを考えることにつながる。

表-3に、閾値を15分、30分、45分、60分とそれぞれとった時の融合推定結果を示す。

全体ではモデルの尤度比は、閾値を大きくとるほど上昇している。これは、閾値以下ではPPデータを使って推定を行っていることから、短時間の滞在の分析では、正確なデータが取れていると思われるPPデータをできるだけ用いた方が精度が向上するためと考えられる。

本研究の結果では、閾値を大きくとってPPを多く使うほど尤度が向上する結果となり、最適な閾値を導き出すまでは至らなかった。評価方法を含めて検討する必要がある。

(3) PT・PP混合比

融合推定においては、PT・PPのデータをどれくらいの割合で用いるかによって、それぞれの重みが変わり、結果に影響が出ると考えられる。これまでの結果では、得られたPTデータ約232万件と、PPデータ23000件を全て用いていたが、ここではPP・PTの混合比に注目し、混合率を1:1, 1:10, 1:101 (PTデータを全て用いた場合) の3つについて、結果の挙動を分析する。なお、1:1, 1:10の混合による分析では、PTデータは全データから当該比率分のデータ数をランダムに抽出した。推定結果は表4に示す通りである。

推定の結果、PTの割合の増加に伴ってサンプル数が増えるため、パラメータの値や尤度比も増加し、パラメータが安定して有意性が増す結果となっている。各パラメータの値の変化をみると、男性ダミーや一部の年齢層のダミー変数では、もともとPPとPTで示している傾向が異なっているため、符号も含めてPTの割合が増すにつれて変化している。

本研究の仮説通り、PTとPPの取得状況の違いが滞在時間の長さのみに依存するとすると、長時間のモデルではPTもPPもパラメータ値が同じになるが、データはそのようにはなっていない。観測欠損の要因は滞在時間以外の交通手段やアクセス・イグレスなどの関係も考えられ、観測の欠損と真値の仮定については今後の課題である。

(4) ゾーン別滞在頻度構成

本推定結果を用いて、PTデータのゾーン別の滞在データをもとに、モデルから求まる滞在の発生確率を計算し滞在活動の頻度を推定した。推定は、日ごとの各個人が各ゾーンへ滞在する確率がプロビット選択モデルより求まることから可能となる。その確率から当該データがどのような頻度で行われる行動であるか計算し、ゾーンごとにデータの集計を行った。分析対象とする横浜都心

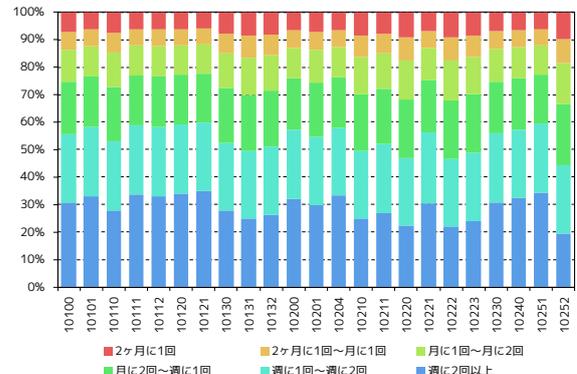


図-8 短時間滞在のゾーン別頻度分布例

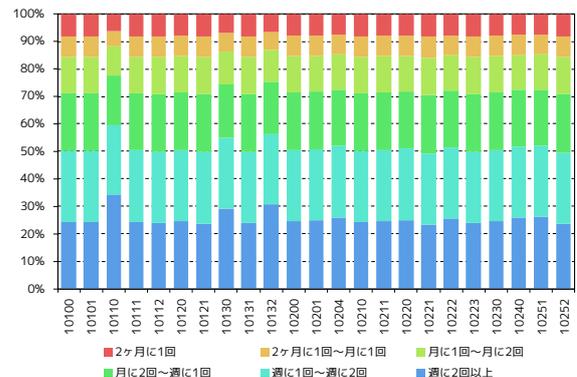


図-9 長時間滞在のゾーン別頻度分布例

部の15の小ゾーンについて、閾値を30分としたときの短時間滞在モデルと長時間滞在モデルから、それぞれ滞在の頻度の構成を推定した結果を図-8、図-9に示す。

図-8と図-9を比べると、短時間滞在では空間ごとの頻度構成比のばらつきがあるが、長時間滞在ではあまり見られない。また、短時間滞在の方が高頻度の割合が大きい。特に高頻度な滞在割合が多いゾーンに注目すると、主要駅から離れ交通の便が比較的に良くない場所が多く、遠くから来る人が少なく、自宅や職場周りでの移動が主となっている地域であると考えられる。

6. 結論・課題

本研究では、PT・PPのデータ融合によるプロビット選択モデルによる滞在分析によって、短時間と長時間の滞りで選択に及ぼす説明変数の影響が異なること、またそれらの影響を有意に表すことができるデータがPTとPPでは取れ方として異なることを示した。推定結果からも、大サンプルのPTデータを用いることで説明力のある結果は得られるが、PPデータと組み合わせることで、自宅や職場周りでの滞在発生効果など、PTだけでは観測精度の問題から評価できなかった要因も考慮することができた。そして、PTデータにデータ融合によ

るモデルを適用して、滞在行動の空間別の特徴の違いを表すことのできる可能性を示した。

本研究の課題として、取得データの選択性バイアスを考慮できていない点が挙げられる。実際の観測では、真値である行動データがわからない中で、PT・PP調査のようなデータでは、真値の行動データから何らかのバイアスが掛かった形で抽出が行われていると考えることができる。真値のデータについてはシミュレーションにより発生させて分析を行う手法が考えられ、Bierlaire et al. (2008)⁶が提案した離散選択モデルにおける実選択肢をもとにした選択肢抽出手法などが参考となる。

また、今回の分析では、空間的な集計の違いや、滞在活動場所のアクセスイグレスの考慮がなされていない。滞在活動場所と来街する際のターミナルとなる駅や駐車場の配置の問題は、都心部での活動を考える上で重要であり、活動需要と交通行動のシミュレーションにつなげる上でも課題である。

謝辞：本研究を実施するに当たって文部科学省科研費基盤 A「プローブ技術を援用したデータフュージョン理論による総合的交通行動調査の高度化（代表：羽藤英

二）」の協力を受けた。ここに感謝の意を表す。

参考文献

- 1) 羽藤英二, 大村朋之, 三谷卓摩: 位置-加速度記録を用いた活動パターンと移動負荷の計測に関する研究, 土木計画学研究・講演集, Vol.45, CDROM, 2012.
- 2) 井坪慎二, 羽藤英二, 中嶋康博, 情報技術の活用による交通行動調査の効率化・高度化に関する研究, 土木計画学研究・講演集, Vol.31, CDROM, 2005.
- 3) 名取義和, 谷下雅義, 鹿島茂: パーソントリップ調査における回答誤差とその発生要因, 土木計画学研究・論文集, Vol. 17, pp. 155-162, 2000.
- 4) Stopher, P., Greaves, S., Missing and Inaccurate Information from Travel Surveys: Pilot Results, Working paper (University of Sydney. Institute of Transport and Logistics Studies), 2010.
- 5) 森川高行, 山田菊子: 系列相関を持つ RP データと SP データを同時に用いた離散型選択モデルの推定法, 土木学会論文集, Vol.476, pp.11-18, 1993.
- 6) Bierlaire, M., Bolduc, D., McFadden, D., The Estimation of Generalized Extreme Value Models from Choice-Based Samples, Transportation Research B, Vol. 42, pp. 381-394, 2008.

(2012.8.3 受付)

Destination Choice Model Using PT/PP Data with Different Observation Scale and Resolution

Sohta ITOH and Eiji HATO

The analysis of generation of activities in urban area is essential for traffic control and spatial designing. As high-resolution and massive data can be obtained, we can analyze traffic behavior by data fusion of these data. In this study, we focus on the difference of Person Trip (PT) data and Probe Person (PP) data. Using PP and PT data, we compose a probit choice model and estimate the model by data fusion. The results show the difference of observation in PT and PP survey and the implication of spatial activity demands.