

統計的に安定な因果推定による 都市間交通需要モデルの改良

片川儀治¹・塚井誠人²

¹正会員 広島市都市整備局都市計画課（〒730-8586 広島市中区国泰寺町一丁目6番34）
E-mail: x2r.p249023@gmail.com

²正会員 広島大学大学院工学研究院（〒739-8527 広島県東広島市鏡山1-4-1）
E-mail: mtukai@hiroshima-u.ac.jp

本研究では、多重共線性が懸念される交通サービス水準（LOS）変数を含む交通需要予測モデルに関して、統計的に安定な因果推定手法を適用したモデル推計の改良効果について検討した。具体的には、傾向スコアによる標本の重みづけによる適合度や有意水準の改善効果を明らかにすることを目的とする。傾向スコアに基づく重みづけで課題となる、サンプルインフレーションの解消を図るMatching Weight法と従来の逆確率（IPW）法、および重みづけなしのOLSモデルの推定結果を比較したところ、Matching Weight法による推定が、最も効果が高いことが明らかとなった。

Key Words : *minimum route frequency, propensity score, gravity model*

1. はじめに

都市間旅客の需要予測ではこれまでに様々な手法が提示されているが、実務的には四段階推定法が依然として重要な位置を占めている。従来の手法に関して、実測値と予測値の乖離が課題として指摘されている。それらの乖離の原因として、予測対象時点までに起こる社会情勢の変化、需要予測モデルの説明変数の不備に加えて、モデルの説明変数として用いた変数間の相関の問題（多重共線性）は無視できない。国土交通省は、交通需要予測モデルの多重共線性の影響に関して、説明変数間に強い相関関係が存在するとパラメータが推定できない、推定パラメータの符号条件が異なるなどの、需要予測の信頼性の低下を招くなどの課題を具体的に挙げて、多重共線性の除去に配慮したモデル推定の重要性を述べている。

交通需要予測モデルでは、説明変数として重要なLOS（交通サービス水準：Level Of Service）変数間に、一定の相関（多重共線性）が見られることが多い。これに対して、説明変数間の相関係数、あるいは推定されたモデルから算出された条件数（VIF）などの指標に基づいて、多重共線性が疑われる変数の組み合わせのうち、一方をモデルから除外する方法が最も一般的である。しかし、この手法では除外変数が交通量に及ぼす影響は評価できないため、多重共線性によるパラメータ推定への影

響を緩和する別のアプローチが必要である。

本研究では、多重共線性を示す複数のアクセス変数、LOS変数を需要予測モデルの説明変数として用いた場合に、それらの因果推定を安定させる手法として、標本の重みを調整する傾向スコア法に着目する。具体的には、傾向スコアを、LOS変数への割り当て確率として、傾向スコアの逆確率（これをIPW推定量という）で標本に重みをかけることによって、多重共線性の緩和を図る。これにより、モデルパラメータの符号や有意水準、ならびに適合度の改善について検証する。

2. 傾向スコアによる標本の重みづけ

傾向スコア法は、無作為割当てが不可能な調査観察研究において因果効果を推定する方法として、Rosenbaum and Rubin¹⁾によって提案された。傾向スコア法により推定される因果効果は、非ランダム化実験において、ある特定の処置を受けた群（処置群）と、処置を受けなかった群（対象群）の結果の差、すなわち処置効果である。処置群と対照群の結果を、それぞれ y_1 、 y_0 とすると、推定される平均処置効果 T は、母集団における期待値を用いて、式(1)のように表せる。

$$T = E(y_1) - E(y_0) \quad (1)$$

非ランダム化実験では、ある処置の処置群と対照群とは体系的に異なるため、結果を直接比較することは難しい。そこで、バランシングスコアを用いて直接的な比較を行う。バランシングスコアは、共変量ベクトルの条件付き分布が、処置群と対照群に対して等しくなるような関数である。傾向スコアは、観測された共変量の分布が、処置群に割当てられる傾向を示す条件付き確率であり、最も粗いバランシングスコアと呼ばれる。

星野²⁾は、因果推論に関して、欠測データを扱う調査観察研究における統計分析手法を整理した。傾向スコアは真値が得られないため、まずデータから推定値を算出して、その値を結果変数モデルに導入する。このように、基本的に傾向スコアを導入した因果効果の推定は、2段階で行う。星野は、Rubinが示した傾向スコアを用いたIPW (Inverse Probability Weighting) 推定量が、ロバストな推定に有用であるとしている。IPW推定量は、傾向スコアの逆数による重み付けによって求める。処置群に対する個人*i*の傾向スコアを e_i とすると、処置群の結果 y_1 と対照群の結果 y_0 の不偏推定量は、IPW推定量を用いて、それぞれ式(2)、式(3)のように表せる。ここで z_i は処置群、対照群を表す2値変数である。

$$\hat{E}(y_1) = \sum_{i=1}^N \frac{z_i y_i}{e_i} \bigg/ \sum_{i=1}^N \frac{z_i}{e_i} \quad (2)$$

$$\hat{E}(y_0) = \sum_{i=1}^N \frac{(1-z_i) y_i}{1-e_i} \bigg/ \sum_{i=1}^N \frac{(1-z_i)}{1-e_i} \quad (3)$$

Rosenbaumらによる傾向スコア法の提案以降、医療や教育の現場で、傾向スコアを適用した介入効果の推定が行われている。

Gumら³⁾は、冠動脈疾患患者に対するアスピリンの有用性の研究を行った。これまでアスピリンが心筋梗塞を減らすことが示唆されてきたが、アスピリンの服用と、他の変数(年齢や他の治療薬の服用、喫煙や心臓疾患に関する検査指標)との多重共線性を調整した研究は無かった。そこで、アスピリン服用者と非服用者の両方を含む、冠動脈疾患患者(またはその疑いのある)6174人に対して調査・分析を実施した。単純な死亡率は、アスピリン服用群、非服用群ともに4.5%であり、有意差がなかった。しかし、アスピリン服用群と非服用群の双方に関係する共変量として、年齢や他の治療薬の服用、喫煙や心臓疾患に関する検査指標など34変数を説明変数として算出した傾向スコアを用いて調整を行ったところ、服用群と非服用群の死亡率はそれぞれ4%、8%となり、有意差が見られた。

中澤⁴⁾は、傾向スコアを用いて出身地域の違いによって、高等教育機関への進学機会に影響があるか否かにつ

いて検討を行った。出身地域を独立変数として、性別、両親の学歴、父親の職業などの共変量を調整して、出身地域が進学率に及ぼす真の因果効果を推定した。その結果、重みの調整後も出身地域の影響は消えることがなく、特に大学と専修学校への進学率の地域格差は拡大するという結果が得られた。

Dinardo, Fortin, and Lemieux⁵⁾は、傾向スコアを応用して、ある事象に影響する要因を分解するDFL法を提案した。山口⁶⁾は、DFL法を用いて、性別が管理職割合へ与える影響を推定した。この研究では、女性が管理職に就いている割合が非常に低いため、性別の効果が正しく推定されない可能性に着目して、女性正社員の教育程度、年齢、勤続年数などの人的資本特性が男性と等しい反事実的状況の下で、性別が管理職割合に与える影響を推定した。推定された傾向スコアモデルの説明変数は人的資本特性、被説明変数は性別であり、通常考えられる因果性とは逆の設定でモデルを推定している。その結果、女性の人的資本特性が男性と同等でも、課長以上管理職の男女格差の21%、係長以上管理職格差の30%しか減少しないことを示した。また、就業時間の男女差を男女間で同等とした場合、それが管理職割合の男女差を付加的に説明できる程度は課長以上で18%、係長以上で13%であることを示した。すなわち、人的資本の男女差が課長以上、係長以上の役職に及ぼす影響40%に過ぎず、残りの60%は男女差が管理職割合に影響していると結論付けている。

傾向スコアによる重み設定に関しては、サンプルインフレーションの問題が指摘されている。これは、式(2)、または式(3)において傾向スコアの推定値 e_i が、それぞれ0または1に極めて近い値をとる標本について、そのIPW重みが非常に大きくなることによって、その標本に与えられる重みが極端に大きな値をとり、他の標本の重みが相対的に小さくなることによって生じる問題である。この課題に対してLi⁷⁾は、式(4)に示す、Matching Weightを提案した。

$$W_1, W_0 = \frac{\min(1-e_i, e_i)}{z_i e_i + (1-z_i)(1-e_i)} \quad (4)$$

式(4)より、Matching Weightは処置群と対照群のそれぞれについて、必ず1よりも小さな値をとる。具体的には、 e_i が処置群への割り当て確率を表すとき、重みが1よりも小さな処置群の標本は、 e_i が0.5を超える場合に限られること/重みが1よりも小さな対照群の標本は、 e_i が0.5を超えない場合に限られることがわかる。以上より式(4)による重みの設定は、サンプルインフレーションを防ぐ働きがある。Liは、Matching Weightによって、処置群と対照群からの期待標本抽出確率の分布と傾向スコアの分布が同等になること、および両者が漸近的に正規分布することを示した。

3. 都市間旅客交通需要分析への適用

(1) 分析の手順

本研究では、旅行者を対象とした Choced based sampling のため、LOS についてランダム割り当てが不可能な調査となる都市間旅客交通需要に関して、LOS が旅客交通需要に及ぼす影響が、傾向スコアによる標本重みづけによる改良効果を検証する。

交通需要と LOS の間に多重共線性が見られる原因として、交通網整備による人口集積の促進と、人口集積による交通網整備という、累積的な因果関係の存在が考えられる。すなわち、交通需要予測や土地利用・交通モデルにおいては、アクセシビリティの空間的な分布と発生・集中交通の空間的な分布の間では、モデルの利用目的に応じて、選択的に因果関係が仮定されることが多く、その結果として LOS 変数間にも多重共線性が生じる。既往研究において DFL 法が提案されていることを踏まえると、推定するモデルを特定化する際に、必ずしも変数間の因果性が適切に特定できていなくても、傾向スコア法による重みづけによってモデルの改良が期待できると考えられる。さらに、標本重みの推定の際には、地理的な特性を反映しやすい LOS 変数を用いたサンプルインフレーションに対処するため、Matching Weight を用いることとする。

以下の分析では、2005 年の幹線旅客純流動データから、平日、および休日の鉄道、航空を代表交通機関とする、全目的の居住地～目的地間の都道府県間 OD 表（ただし北海道は 4 地域区分、鉄道の無い沖縄を除く 49 ゾーン間）に基づいて、重力モデルの推定を行う。LOS 水準としては、同年の時刻表から最短経路探索によって独自に推計した代表都市（ゾーン内の人口最大／または都道府県庁所在地）間の所要時間、経路上に現れる 3 時間あたりの最低運行頻度、経路距離などを用いる。

Matching Weight を適用した重力モデルの推定手順を示す。以下では LOS 変数のうち、最低運行頻度と、所要時間を経路距離で除した経路平均速度に着目して、これらを傾向スコア算出基準として用いる。具体的には、標本を LOS の高低に対応して 1/2 に分割するため、鉄道または航空の最低運行頻度／経路平均速度を全 OD ペアに対して昇順に並べて中央値を求め、OD ペアを 2 分割した。この指標を傾向スコア分析における割り当て変数とみなして、割り当て確率を表す 2 項ロジットモデルを推定する。推定されたロジットモデルから Matching Weight を算出して、重みづけ最小二乗法によって、重力モデルを推定する。

モデル推定に用いる標本数は、各ゾーン内々と 3 大都市圏内々を除いた 2322 サンプルである。

(2) 傾向スコアモデルと重力モデルの推定結果

分析の結果、経路平均速度を割り当て変数とした場合はモデルの改良効果が見られなかったため、以下では航空経路の最低運行頻度を割り当て変数とした場合の推定結果を示す。なお本分析では、航空経路の最低運行頻度が低いグループに 1 を、高いグループに 0 を設定して傾向スコアモデルを推定した。

表 1 傾向スコアモデル（航空最低運行頻度）の推定結果

説明変数	推定値	T値
発地生産年齢人口	-0.032 **	-6.03
航空速度	0.170 *	2.56
鉄道速度	-0.019	-0.11
鉄道最低運行頻度	-0.023 **	-3.15
定数	0.451 *	2.11
最終尤度	-545.910	
尤度比	0.661	
サンプル	2322	

***: 0.1% 有意, **: 1% 有意, *: 5% 有意

表 1 に、航空最低運行頻度に関する傾向スコアモデルの推定結果を示す。モデルの適合度は比較的高く、良好なモデルが推定できている。表 2、および表 3 に、平日および休日の重力モデルの推定結果を、それぞれ示す。両表より、モデルの自由度調整済決定係数は、それぞれ 0.75 程、0.73 度の値を示しており、適合度は中程度である。OLS モデルと比較すると、Matching Weight モデルの決定係数は、平日・休日とも約 0.01 程度改善していることがわかる。一方、IPW モデルの適合度は OLS モデルと同程度であり、適合度の改善効果は見られない。パラメータの符号に着目すると、3 モデルの符号は全て一致している。なお両モデルの鉄道最低頻度に関しては、正の符号が期待されるものの、パラメータは負号で推定されており、条件を満たさない。しかし休日モデルに関しては、Matching Weight、IPW の両モデルで有意ではない結果を得ており、妥当な結果となっている。さらに、割り当て変数として設定した航空最低運行頻度に関しては、特に休日で有意水準が改善している。

一方で、鉄道の経路速度は全モデルで負号を示しており、期待された条件を満たさない。これは、近距離の需要の多い OD の平均運行速度が遅いことを反映していると考えられるが、少なくとも傾向スコアの導入によって符号が改善する効果は見られないことが明らかになった。

4. おわりに

傾向スコアモデルの導入による、重力モデルの適合度の改善効果について検討を行った。その結果、Matching Weight 法の有効性が明らかとなった。

表1 都道府県間旅客需要モデルの推定結果(平日)

平日 説明変数	Matching Weight		IPW		Not-weighted (OLS)	
	推定値	T値	推定値	T値	推定値	T値
発サービス従業	2.377 ***	17.47	2.365 ***	16.18	2.337 ***	15.96
着サービス従業	0.290 *	2.12	0.280 *	2.01	0.331 *	2.26
発従業/企業	-1.066 ***	-7.04	-1.060 ***	-6.56	-1.023 ***	-6.35
着従業/企業	0.965 ***	6.40	0.974 ***	6.34	0.893 ***	5.53
航空時間	-0.134	-1.57	-0.126	-1.43	-0.076	-0.85
鉄道時間	-0.861 ***	-24.22	-0.874 ***	-24.27	-0.833 ***	-23.34
航空経路速度	0.553 ***	8.30	0.587 ***	8.70	0.564 ***	8.23
鉄道経路速度	-0.737 ***	-7.74	-0.738 ***	-7.51	-0.684 ***	-7.01
航空最低頻度	0.096 *	2.43	0.093 *	2.29	0.045 *	1.98
鉄道最低頻度	-0.029	-0.74	-0.032	-0.81	-0.052	-1.35
定数	-23.687 ***	-26.69	-23.684 ***	-25.75	-24.232 ***	-25.60
残差標準偏差	0.8977		0.9160		0.9219	
決定係数	0.7660		0.7533		0.7533	
調整済	0.7649		0.7523		0.7522	
標本数			2322			

***:0.1%有意, **:1%有意, *:5%有意

表2 都道府県間旅客需要モデルの推定結果(休日)

休日 説明変数	Matching Weight		IPW		Not-weighted (OLS)	
	推定値	T値	推定値	T値	推定値	T値
発サービス従業	2.369 ***	16.51	2.372 ***	15.44	2.372 ***	15.39
着サービス従業	0.354 *	2.46	0.351 *	2.39	0.336 *	2.17
発従業/企業	-1.129 ***	-7.07	-1.137 ***	-6.70	-1.132 ***	-6.67
着従業/企業	0.889 ***	5.59	0.895 ***	5.54	0.900 ***	5.30
航空時間	-0.240 **	-2.67	-0.220 *	-2.37	-0.212 *	-2.26
鉄道時間	-0.866 ***	-23.10	-0.882 ***	-23.31	-0.841 ***	-22.37
航空経路速度	0.659 ***	9.37	0.695 ***	9.80	0.610 ***	8.47
鉄道経路速度	-0.793 ***	-7.89	-0.799 ***	-7.74	-0.715 ***	-6.95
航空最低頻度	0.134 **	3.21	0.139 **	3.27	0.049 *	2.04
鉄道最低頻度	-0.064	-1.57	-0.068	-1.62	-0.093 *	-2.27
定数	-22.417 ***	-23.93	-22.555 ***	-23.33	-22.772 ***	-22.85
残差標準偏差	0.9473		0.9628		0.9708	
決定係数	0.7371		0.7264		0.7265	
調整済	0.7359		0.7252		0.7253	
標本数			2322			

***:0.1%有意, **:1%有意, *:5%有意

参考文献

- 1) Rosenbaum, R., Rubin, B. : The central roles of the propensity score in observational studies for causal effects, *Biometrika*, Vol.70, pp.41-55, 1983
- 2) 星野崇宏: 調査観察データの統計科学—因果推論・選択バイアス・データ融合, 岩波書店, 2009
- 3) Gum, P., Thamillasan, M., Watanabe, J., Blackstone, E. and Lauer, E. : Aspirin use and all-cause mortality among patients being evaluated for known or suspected coronary artery disease, *JAMA*2001, pp.1187-1194, 2001
- 4) 中澤渉: 高校進学機会の地域間不平等, 東洋大学社会学部紀要, vol.48-2, pp.5-18, 2010

- 5) Dinardo J, Fortin N, and Lemieux T : Labor Market Institutions and the Distribution of Wages,1973-1992:A Semi-parametric Approach, *Econometrica*, vol.64, pp.1001-1044, 1996
- 6) 山口一男: ホワイトカラー正社員の管理職割合の男女格差の決定要因—女性であることの不当な社会的不利益と、その解消施策について, *RIETI Discussion Paper Series 13-J-069*, 2013
- 7) Li, L. : Propensity Score Analysis with Matching Weights, *Collection of biostatistics research archive*, paper 79, 2011.

(2014.04.25 受付)