

Mass Point手法の課題に関する検討*

Investigating the Problems of Mass Point Approach*

杉恵頼寧**・張 峻屹***・藤原章正**・宮地武志****

By Yoriyasu SUGIE **, Junyi ZHANG***, Akimasa FUJIWARA ** and Takeshi MIYAJI****

1. はじめに

個人の交通行動の違いをモデルの中でいかに考慮するかは非常に重要な問題である。今まで、母集団全体に共通のモデル構造（例えば、共通の説明変数パラメータや共通の誤差項など）を仮定するのが一般的であったが、最近になって、マーケット・セグメンテーション手法のように、母集団をグループ分けしてそれぞれ独立なモデルを構築する方法が次第に実用化されつつある。マーケット・セグメンテーション手法は個人の観測可能な特性（性別、年齢、収入など）を基準に母集団を分割する。

これに対して、今まで仮定した誤差項の個人間における均質性（個人間で共通の誤差項をもつこと）を改め、個人によって異なると仮定したMass Point（略してMP）手法が最近適用されるようになった。個人によって異なる誤差項はここで特に非観測異質性と呼ぶ。この非観測異質性は潜在的に存在する個人の性格、態度、動機、嗜好などのような測定が難しい特性の影響を考慮することができる。しかし、（1）MPをどう解釈するか、（2）非観測異質性パラメータが時間的に安定するかどうか、（3）非観測異質性パラメータを母集団のパラメータとして用いるのが適切かどうかが課題として残されている。

そこで、本研究は以下のように構成される。第2節ではMP手法を概説する。第3節ではMPの意味を実証的に解釈する。第4節ではパネルデータから推定した非観測異質性パラメータの時間的安定性とその母集

団代表性問題を検討する。利用するデータは当研究室が実施した広島市の新交通システムの利用意識に関する同じ時間間隔の3時点パネル調査データ（1987、1990、1993）である。

2. Mass Point手法の概要

本研究はMP手法を中心としたもので、ここではその概要を簡単に説明しておく。

MP手法はLaird (1978)¹⁾とLindsay (1983a, b)^{2,3)}により確立されたもので、特にLindsay手法では簡単な仮説を立てており、いろいろな関数形に拡張しやすい。まず、以下の確率密度関数を考える。

個人 i に関する変数 y_i の確率密度関数を $g_i(y_i)$ 、その非観測変数 δ_i の確率密度関数を $f(\delta_i)$ 、 δ_i を条件とする y_i の確率密度関数を $h_i(y_i | \delta_i)$ とする。

$$g_i(y_i) = \int_{\delta_i} h_i(y_i | \delta_i) f(\delta_i) d\delta_i \quad (1)$$

式(1)において、 $h_i(y_i | \delta_i)$ のみが既知である。

Laird (1978)とLindsay (1983a, b)は $h_i(y_i | \delta_i)$ がいくつかの条件を満たせば、式(1)を式(2)のように離散的に表現できることを証明した。

$$g_i(y_i) = \sum_{k=1}^K h_i(y_i | \xi_k) p_k \quad (2)$$

そして、以下の制約条件が存在する。

$$0 \leq p_k \leq 1, \sum_{k=1}^K p_k = 1 \quad (3)$$

ただし、 ξ_k, p_k はそれぞれ k 番目のMPの位置と重みを表すパラメータである。

Laird (1978)が仮定した条件は比較的厳しいので、適用範囲が狭いに対して、Lindsay (1983a,b)の条件は

* キーワード：交通行動分析、手段選択

** 正会員、工博、広島大学大学院国際協力研究科
(東広島市鏡山1-4-1、TEL & FAX: 0824-24-7826)

*** 正会員、工博、広島大学工学部第四類建設系
(東広島市鏡山1-4-1、TEL & FAX: 0824-24-7849)

**** 学生員、広島大学大学院国際協力研究科
(東広島市鏡山1-4-1、TEL & FAX: 0824-24-7849)

一般性をもつため、適用範囲が広い。ここでは、Lindsay(1983a,b)の条件のみを示す。

条件 (1) : $h_i(y_i | \delta_i)$ は有限関数 (bounded function)

条件 (2) : $f(\delta_i)$ の分布関数 $\int_{\delta_i} f(\delta_i) d\delta_i$ は右連続 (right continuous)、非減少関数

このように、Mass Pointとは非観測異質性の連続分布を離散化した点のことである。

3. Mass Pointの解釈

今まで、著者等^{4,5,6}が交通機関選択を対象にパネルデータを用いたMP手法の有効性を実証してきた。しかし、冒頭で述べたように、MPが何を意味するかは十分説明できなかった。これを検討するために広島市の新交通システムの利用意識に関する3時点パネルデータを用いる。対象交通手段が新交通システム(NTS)、自家用車(CAR)とバス(BUS)である。調査概要を図1に示す。

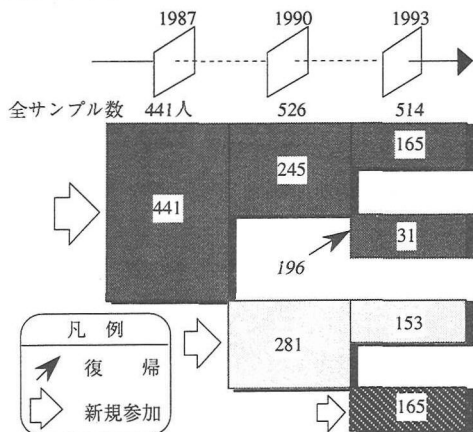


図1 新交通システム3時点パネルデータの概要

この節では図1に示される165人の3時点パネルデータのみを利用する。

パネルデータを用いて、多項交通機関選択モデルの中に非観測異質性パラメータ δ を導入すれば、以下の条件付き選択確率 $P_i(j|\delta)$ が得られる。

$$P_i(j|\delta) = \frac{\prod_{t=1}^{T_i} \prod_{j=2}^J [\exp(\beta'(x_{ijt} - x_{i1t}) + \delta_{jk})]^{\omega_{ijt}}}{1 + \sum_{j=2}^J \exp(\beta'(x_{ijt} - x_{i1t}) + \delta_{jk})} \quad (4)$$

ただし、 x_{ijt} は時間 t における個人 i 、選択肢 j の

説明変数ベクトル、 β はそのパラメータである。 x_{i1t} はある基準選択肢の説明変数で、本研究ではそれを自家用車とする。 T_i は個人 i の参加時点数、 ω_{ijt} は0か1の値をとる選択肢ダミー変数である。

$P_i(j|\delta)$ は第2節で示したLindsayの条件を満たすために、MP手法に基づいた個人 i の同時生起確率 p_i が以下のように表される。

$$p_i = \sum_{k=1}^K \left\{ \frac{\prod_{t=1}^{T_i} \prod_{j=2}^J [\exp(\beta'(x_{ijt} - x_{i1t}) + \xi_{jk})]^{\omega_{ijt}}}{1 + \sum_{j=2}^J \exp(\beta'(x_{ijt} - x_{i1t}) + \xi_{jk})} \right\} p_k \quad (5)$$

式(5)を推定するにあたり、用いる説明変数はアクセス時間(分)、乗車時間(分)、コスト(100円)、待ち時間(分)の4変数である。推定結果を表1に示す。同表から、説明変数パラメータの推定値の変動と最大尤度の変化傾向から判断すると、収束MP数が2であることが分かる。つまり、母集団には2つの同質な潜在グループが存在することを意味する。

表1の推定結果を用いて、潜在的グループごとの各時点における交通機関別の推定選択確率を計算し、紙面上の制限で1987年の結果のみを図2に示す。

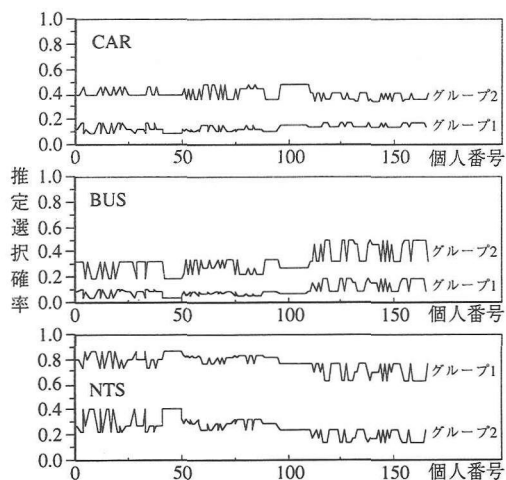


図2 1987年のグループごとの推定選択確率

図2から、グループ1では新交通システムを選択する確率が高く、グループ2では自動車とバスを選択する確率が高い。他の時点においても同様な結果が得られた。つまり、グループ1は新交通システムを好む個人の集合で、グループ2は新交通システムを好まな

表1 3時点パネルデータを用いた交通機関選好モデルの推定結果

説明変数	MP = 1	MP = 2	MP = 3	MP = 4	MP = 5
アクセス (分)	-0.064*	-0.064	-0.064	-0.065	-0.062
乗車時間 (分)	0.002	0.005	0.005	0.006	0.005
コスト (100円)	-0.018	-0.027	-0.027	-0.023	-0.024
待ち時間 (分)	-0.102**	-0.113**	-0.113**	-0.106**	-0.112**
ρ_1^*		0.741**	0.738**	0.753**	0.300**
ξ_{12}	0.875**	0.818**	0.819*	0.553	0.875*
ξ_{13}	1.896**	3.048**	3.054**	2.907**	3.135**
ρ_2^*			0.435*	0.586**	0.419**
ξ_{22}		1.039**	0.997*	0.775*	0.927*
ξ_{23}		0.771**	0.499	0.493	-0.310
ρ_3^*				0.533	-0.132
ξ_{32}			1.074*	0.771	-0.614
ξ_{33}			0.980	1.366	5.927**
ρ_4^*					0.587**
ξ_{42}				-0.263	0.682
ξ_{43}				1.154	3.085**
ρ_5^*					1.108**
ξ_{52}					1.594**
ξ_{53}					
初期尤度	-543.81	-543.81	-543.81	-543.81	-543.81
最大尤度	-499.61	-482.96	-482.96	-482.34	-482.99
尤度比	0.076	0.104	0.101	0.099	0.095
個人数	165	165	165	165	165

($\rho_k = \rho_k^* \times \rho_k^*$; ξ_{k2} , ξ_{k3} はそれぞれBUSとNTSの位置パラメータで、自動車の ξ を0と仮定した。

MP: Mass Point; *, 5%; **, 1%で有意)

いものである。ここでいうグループとは母集団全体からみた表現であり、個人からみるとどのグループにも属する可能性があり、その確率を p_k として与える。実際、どのような要因がこのような傾向をもたらしたかを分析する必要がある。

この節では3時点パネルデータのみに対して、MPの意味を簡単に解釈したが、この解釈は母集団に拡張できるかどうかに関しては検討していない。そこで、次節ではこのパネルデータから推定した非観測異質性パラメータ ξ_{ik} , p_k の時間的安定性とその母集団の代表性問題を検討する。

4. 非観測異質性パラメータの時間的安定性とその母集団の代表性問題

まず、時間的安定性を分析するために第3節の165人3時点パネルデータの推定結果を、3時点目のデータを除いた他の2時点（1987、1990）パネルデータの推定結果と比較する。後者の推定結果を図3に示す。同図から収束MP数が3時点パネルデータの推定結果と同様に2であり、収束MP数が時間的に安定すると

言える。そして、両パネルデータの推定パラメータ値に差があるか否かに関するt検定の結果を表2に示す。説明変数の推定パラメータと同様に ξ_{ik} , p_k の推定値にも95%の有意水準で差がないことが分かる。

表2 2、3時点パネル間の推定パラメータのt検定

アクセス	乗車時間	コスト	待ち時間	ρ_1^*
0.515	1.19	1.009	0.043	0.622
ξ_{12}^*	ξ_{13}^*	ξ_{22}^*	ξ_{23}^*	
0.658	0.067	0.637	0.241	

つぎに、 ξ_{ik} , p_k の母集団の代表性問題を検証する。表3に図1の全サンプルを用いた交通機関選好モデルの推定結果を示す。この結果を母集団のものとする。165人の3時点パネルデータを用いたモデルでは収束MP数が2であるのに対して、母集団のモデルでは3になっている。しかも、各 ξ_{ik} , p_k はほとんど有意であり、パネルサンプルと母集団との間に差があることを意味する。つまり、母集団には3つの潜在的なグループが存在するが、パネルデータではそれをすべて説明することができていない。言い換えれば、パネルデータに残るサンプルが偏ったものであり、母集団

表3 図1の全サンプルを用いた交通機関選択モデルの推定結果

説明変数	MP = 1	MP = 2	MP = 3	MP = 4	MP = 5
アクセス (分)	-0.072**	-0.068**	-0.080**	-0.082**	-0.082**
乗車時間 (分)	-0.003	-0.002	-0.001	-0.001	-0.001
コスト (100円)	-0.077**	-0.079**	-0.058**	-0.086**	-0.085**
待ち時間 (分)	-0.069**	-0.083**	-0.080**	0.079**	-0.077**
ρ^*_1		0.509**	0.591**	0.551**	0.541*
$\hat{\epsilon}_{12}$	0.905**	-0.629**	0.593	0.737	0.786
$\hat{\epsilon}_{13}$	1.886**	3.230**	2.915**	2.789**	0.588*
ρ^*_2			0.597**	0.580**	0.713
$\hat{\epsilon}_{22}$		1.094**	1.014**	0.779	2.712**
$\hat{\epsilon}_{23}$		1.477**	2.439**	2.747**	0.733**
ρ^*_3				0.631**	1.016**
$\hat{\epsilon}_{32}$			1.163**	1.080**	0.807
$\hat{\epsilon}_{33}$			0.485	0.727	-0.346
ρ^*_4					0.457
$\hat{\epsilon}_{42}$				0.123	1.304
$\hat{\epsilon}_{43}$				0.626	
ρ^*_5					0.555
$\hat{\epsilon}_{52}$					1.093
$\hat{\epsilon}_{53}$					-0.077**
初期尤度	-1625.90	-1625.90	-1625.90	-1625.90	-1625.90
最終尤度	-1496.40	-1481.35	-1477.98	-1477.85	-1477.84
尤度比	0.078	0.086	0.087	0.086	0.086
個人数	887	887	887	887	887

($\rho_k = \rho^*_k \times \rho^*_k$; $\hat{\epsilon}_{k2}$, $\hat{\epsilon}_{k3}$ はそれぞれBUSとNTSの位置パラメータで、自動車の $\hat{\epsilon}$ を0と仮定した。MP: Mass Point; *: 5%, **: 1%で有意)

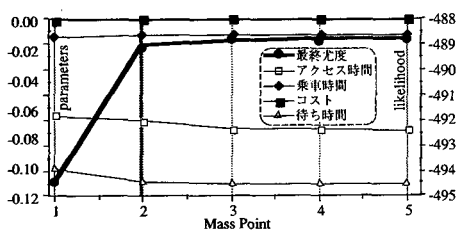


図3 165人3時点パネルデータのうちの1987、1990年の2時点のデータを用いた推定結果

全体の行動を説明するために純粋なパネルデータ以外のサンプル情報活用する必要がある。

5. おわりに

本研究では今まで残されたMP手法の課題であるMPの意味合い、パネルデータから推定した非観測異質性パラメータの時間的安定性を明らかにした。そして、パネルデータから推定した非観測異質性パラメータをそのまま母集団パラメータとして利用できないことを指摘し、全調査情報を活用する必要性を明らかにした。しかし、どのような要因が母集団やパネルサンプルの潜在的グループ分けに影響するか

を明らかにしていない。特に、非観測異質性パラメータの時間的安定性と母集団の代表性問題に関してはさらに時点数の多いパネルデータを用いて深く検討する必要がある。

参考文献

- 1) Laird N.: Nonparametric maximum likelihood estimation of a mixing distribution, Journal of American Statistical Association, Vol.73, No.364, pp.805-811, 1978.
- 2) Lindsay B.G.: The geometry of mixture likelihoods, Part I, Annals of Statistics, Vol.11, pp.86-94, 1983a.
- 3) Lindsay B.G.: The geometry of mixture likelihoods, Part II, Annals of Statistics, Vol.11, pp.783-792, 1983b.
- 4) 杉恵頼寧・張峻屹・藤原章正: 個人の異質性による交通機関選択モデルの構造分析、土木計画学研究・論文集、No.12、pp.425-434, 1995.
- 5) 杉恵頼寧・張峻屹・藤原章正: 非観測異質性を考慮した多項交通機関選択のMass Pointモデル、第15回交通工学研究発表会論文報告集, pp.217-220, 1995.
- 6) 杉恵頼寧・張峻屹・藤原章正: Mass Point手法により非観測異質性を考慮した交通機関選択のダイナミックモデル、土木計画学研究・論文集、No.13, 1996 (印刷中)。