

IV-345

SPデータに内在するバイアスを考慮したパラメータ推計

—Mass Point Model による—

山梨大学工学部	正員	西井和夫
○山梨県	正員	弦間重彦
京都大学工学部	正員	北村隆一

1. 研究の目的

近年SPデータは、交通行動分析の分野においてしばしば用いられている。しかし、このデータは、同一被験者に対して繰り返し調査の形態をとる場合が多いために、これまでの断面調査での無作為抽出方法と異なり、その有効性とは別に、データに内在する特有のバイアスが含まれている。本研究の目的は、非集計選択モデルを構築する際に、こうしたバイアスを考慮したパラメータ推計法の提案にある。

今、筆者らによってなされているパネル調査データのよう choice-based sampling形式のSPデータを含む場合には、choice-based samplingによる被験者の偏りに起因するバイアスおよびSPデータとして1個人に繰り返し質問するために個人が仮想的に倍増することによるバイアス（以下、意向バイアスと呼ぶ）を考慮したパラメータ推計法を紹介していきたい。

2. 意向バイアスを考慮したパラメータ推計法

Mass Point Model

まず、通常の2項選択におけるSPモデルの尤度関数は、(2-1)式のように示され、この対数尤度を最大にするパラメータを推計する。

$$L^* = \prod_{n=1}^N \prod_{t=1}^T \left(\frac{1}{1 + \exp(\beta_k x_{nt})} \right)^{\delta_{1nt}} \left(\frac{\exp(\beta_k x_{nt})}{1 + \exp(\beta_k x_{nt})} \right)^{\delta_{2nt}} \dots (2-1)$$

β : パラメータベクトル

T : 質問するケースの数

δ_{1nt} : 代替案1を選択=1

代替案2を選択=0

δ_{2nt} : 代替案1を選択=0

代替案2を選択=1

x_{nt} : 個人nのt番目の説明変数値

ところが、このままでは、結局全サンプル数としての $N \times T$ サンプルの個人が倍増された形で推計を行なうことになる。またこのモデルでは、集団全体の各選択肢への意向の偏りを考慮することは不可能である。

そこで、集団全体の各選択肢に対する意向の重み p をパラメータ α によって決定する(2-2)式のように定義する。

$$p = \frac{1}{1 + \exp(\alpha)} \dots (2-2)$$

今、 p および $(1-p)$ を考慮した効用関数に対して、固有ダミー定数 (Mass Point) として、それぞれ m_1 , m_2 を導入することにより、(2-1)式を(2-3)式のように分割できる。

$$L^* = \prod_{n=1}^N \prod_{t=1}^T \left(\frac{1}{1 + \exp(\beta_k x_{nt} + m_1)} \right)^{\delta_{1nt}} \left(\frac{\exp(\beta_k x_{nt} + m_2)}{1 + \exp(\beta_k x_{nt} + m_2)} \right)^{\delta_{2nt}} \\ + (1-p) \prod_{t=1}^T \left(\frac{1}{1 + \exp(\beta_k x_{nt} + m_2)} \right)^{\delta_{1nt}} \left(\frac{\exp(\beta_k x_{nt} + m_2)}{1 + \exp(\beta_k x_{nt} + m_2)} \right)^{\delta_{2nt}} \dots (2-3)$$

$$p = \frac{1}{1 + \exp(\alpha)}$$

α : パラメータ

m_1 : Mass Point 1

m_2 : Mass Point 2

このように定式化することで、集団全体の選択肢に対する偏りを考慮することができ、同時に本来のサンプル数としてパラメータ推計を行なうことが可能となる。

このMass Point Modelは、 $\alpha=0$ で $p=0.5, m_1=m_2=0$ とすると、当然のことながら通常のLogit Modelとなる。もしこのような結果となった場合は、バイアスは存在しないと判断できる。

詳しく説明するために図-1に m_1, m_2, p と効用との関係を示した。これは、ある選択肢(例えば選択肢2)に対してプラス方向(効用を上げる)に m_1 だけシフトした位置で尤度関数へのウェイトが p だけかかり、その選択肢に対してマイナス方向(効用を下げる)には、 m_2 だけシフトした位置で $(1-p)$ のウェイトが、かかることを意味している。すなわち、対象全体が p のウェイトで m_1 分効用を上げる。

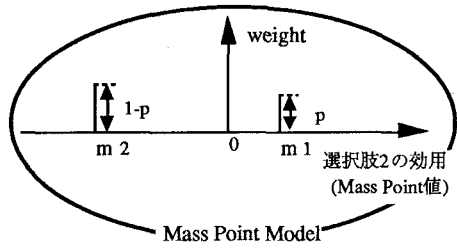


図-1 Mass Point Modelの考え方

一方で $(1-p)$ のウェイトで m_2 分だけ効用を下げるような選択を行っているという考えである。

3. データの適用

本研究の対象データは、4回のパネル調査の中でSPデータの収集を行ったwave2時点(1990年)の当日SC来訪者の323ケースを対象とした。

3-1. SP調査の概要

本節で取り扱うモデルは、SPデータによる買物場所選択モデルである。このSPモデルの説明変数は、政策変数として、「駐車待ち時間」、「駐車料金」、「大型スーパーの有無」、「百貨店の有無」の4変数とし、この政策変数により8ケースの質問を抽出し、SCと中心街の2項選択データを得た。

3-2. パラメータ推計結果

この通常のモデルおよびMass Point Modelによるパラメータ推計結果を表4-1に示す。通常のモデルは、尤度比 (ρ^2 値) が0.265とまずまずの結果が得られた。しかし、ここで注意しなければならないことは、前述のようにこの推計モデルではサンプル数を質問ケース分だけ倍増した1240ケースで推計していること、また集団の各選択肢に対する選好の偏りを考慮していないことである。

そこでMass Point Modelでの推計結果を見てみると、 m_1 の t 値は0.827と弱い規定力であるが、 m_2 のそれは-3.885と大きな規定力を持っており、十分有意なパラメータとなっている。また、 α に対する t 値は0.340と小さな値を示しているが、この α より p および θ を求めると、 $p=0.4753$ となる。このとき、パラメータ θ を、すべて0とすると、選択の偏りは、図4-1のように示すことができる。これより m_1 、 m_2 は中心街の固有ダミー定数なので、符号が負となる m_2 は、中心街の効用をそのパラメータの絶対値 ($|m_2| = |-2.2328| = 2.2328$) だけ低下させる作用をもつ。ゆえに $(1-p)=0.5247$ のウェイトで集団全体が2.2328だけ中心街を過小評価していると言え、 α の t 値は小さいが結果となっているが、全体としてある程度意向バイアスを評価できたこ

とになる。

また同一個人が質問ケース分だけ倍増してしまう問題についても、通常のSPモデルと比べて、本モデルは、明らかに個人・世帯属性に対する t 値が低下していることがわかる。これより、Mass Point Modelでは、このバイアスを考慮したパラメータ推計が行なわれていると言える。

全体の適合度を示す尤度比 (ρ^2 値) を通常のモデルと比べると、0.329とMass Point Modelの方が良好である。やはり同一個人を質問ケース分倍増して推計せねばならない通常の推計方法より、このMass Point Modelの方が現実さに即した考え方であることが実証的にも明らかにされたと言えよう。

表-1 パラメータ推計結果

		Normal Logit		Mass Point	
説明変数		θ	t	θ	t
個人属性変数					
1.ライフスタイル	S	0.878	5.84	0.931	2.67
2.アパレル面積	G	0.460	4.61	0.362	1.57
3.買物形態	S	-0.613	-3.85	-1.152	-3.79
4.滞在時間	S	-0.619	-5.70	-1.256	-4.82
政策変数					
5.駐車待ち時間	D	-7.752	-11.12	-10.203	-11.89
6.駐車料金	D	-3.487	-10.05	-4.708	-11.03
7.大型スーパー	D	1.279	6.31	1.882	6.83
8.百貨店	S	0.670	4.23	0.880	4.81
9. α	—	—	—	0.099	0.34
10. m_1	D	—	—	0.395	0.83
11. m_2	D	—	—	-2.233	-3.89
L (0)		-859.503		-859.503	
L (θ)		-631.735		-576.610	
$\rho^2 (=1-L(\theta)/L(0))$		0.265		0.329	
ケース		1240(155×8)		155	

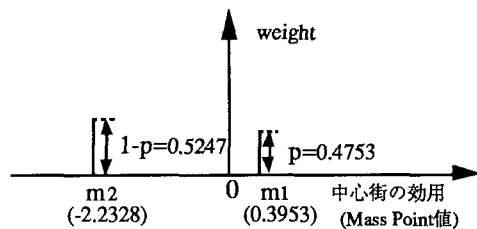


図-2 Mass Point Modelの選択の偏り

4. おわりに

本研究では、データに内在するバイアスを考慮したモデルとして、Mass Point Modelを提案した。データ適用にあたっては、SC側への選択の偏りを考慮して推計を行なうことが可能となった。

なお、この問題については選好の偏りを正規分布に従う確率変数で表したMixing Distribution Modelを併せて検討しており、機会を改めて発表する予定である。