

個人データと集計データとの統合 利用によるモデル構築方法

ESTIMATION OF DISAGGREGATE MODEL USING ADDITIONAL AGGREGATE DATA

森地 茂,* 屋井 鉄雄,** 平井 節生***

By Shigeru MORICHI, Tetsuo YAI and Setsuo HIRAI

The purpose of this paper is to develop the methods to estimate the parameters of behavioral models with disaggregate and aggregate data. Bayesian approach is utilized for combining aggregate data such as passenger counts with the models. The Posterior distribution of the parameters is obtained using parameters only with disaggregate data and likelihood functions which are composed of aggregate data.

Properties of the estimators which are proposed in this study are analyzed with the simulation, and the estimators are compared with alternative estimation methods. The results indicate that the new techniques are so effective to build behavioral models and their features are well understood.

1. はじめに

非集計行動モデルの構築には比較的少数の個人データがあれば良いとされ、その利点故に従来より数多くの現象及び対象のモデル化が行われてきた。但し、サンプル数が少なくなればなるほど、パラメータがより不安定になることは明らかである。これに対し、必要なサンプル数に関する議論が様々に展開されており、幾つもの有益な知見が蓄積されているが、必要十分なサンプルサイズの決定方式については未だ一般論が提案されてはいない状況にある。

一方、観測交通量や大規模調査のデータの様に、集計量としての信頼性が比較的高い統計が、様々な交通現象を対象として収集されている。このようなデータ（本研究ではこれらを集計データと称する。）は、モデル作成上の利用価値が高いと考えられるが、

それだけではモデルパラメータの推定を行えないことが多い。

この時、もし、これら両者の情報を有効に用いてモデル構築が行えるならば、交通需要予測上非常に有意義と考える。何故なら、数少ないが詳細な情報である個人データと、量としての信頼性が比較的高い集計データとの両者を、相互に補完させてモデル構築に活用できるからである。その結果、先に述べた必要サンプル数問題についても、新たな積極的展開が可能になるう。

本研究は以上に述べた点に留意して、以下に示す3つの目的の基で分析検討を行ったものである。

- 1) 非集計行動モデルのパラメータ推定段階に、集計データの持つ情報を取り入れる、一般的な方法を開発すること。
- 2) 開発した方法より求められるパラメータの特性をシミュレーション分析を通して把握すること。
- 3) 開発した方法と既存の方法との関連を調べ、位

* 正会員 工博 東京工業大学教授 工学部土木工学科
(〒152 目黒区大岡山2-12-1)

** 正会員 工博 東京工業大学助手 同上

*** 正会員 工修 建設省

置づけを明確にすること。

2. 従来の研究

非集計行動モデルのパラメータ推定において、複数のデータを用いる試みは、従来より余りなされていない。同モデルは個人の選択行動を表現するためのモデルであり、集計されたデータの利用とは本来的に馴染まないと考えられがちであるが、ミクロ行動とマクロ現象との構造的関連を直接モデル化して行く試みとは別に、1.に述べたような理由から、パラメータ推定としての利用価値は大きいと考えられる。

森地・屋井(1985)では、非集計行動モデルを用いて分布交通量を推計する際に、他のデータソースから得られる集計データを、定数項修正のために活用する方法を提案して、その有効性を実証的な分析より確認している。

また、Quevedo(1985)は、両者のデータを同時に1つの尤度関数で表現して、パラメータを推定する方法を開発している。そこでは、個人の選択が連続量で表され、それが正規分布に従うと仮定しており、一般的な非集計行動モデルに適用可能な方法とはなっていない。

本研究では、以上に示した研究を踏まえた上で、既存統計等の観測交通量や集計データを、より一般的に活用するための統計的推定方法の開発を行っている。

3. 集計データを取り入れた推定方法

本章では、先に述べたパラメータ推定方法の提案を行う。それは、集計データと個人データとの両者の信頼性を考慮して、ベイズ推定法の応用によって合理的にパラメータを推定する方法である。観測交通量等の集計データは各個人の選択結果の総数である。したがって、観測値の信頼性を考慮した上で、モデルから推計される集計量を観測値に近づけるように、パラメータを定めることが望ましいと考えた。

ベイズ推定法は、ベイズの定理をモデルパラメータの推定問題に適用したものである。従って、パラメータの事前分布と、事後的に得られるデータの尤度関数とが与えられれば、両者を用いてパラメータ

の事後分布を導出でき、新たなパラメータを求められる。

本研究では事前分布として、個人データをもとに推定された非集計行動モデルのパラメータ分布を用い、尤度関数として、与えられた集計量データと非集計行動モデルを用いた集計値(モデルパラメータの関数)との差の分布を用いている。前者は最尤推定量であるため、漸近的に正規分布に従うと仮定でき、また後者は正規分布に従うと仮定する。ここで、事後分布は、ベイズの定理より事前分布と尤度関数とを乗じたものに比例するので、その最大値を求めることによって新たなパラメータが求められる。パラメータの解法には直接最大化する方法と、近似解を簡便に求める方法とを考えた。それらを以下順に解説する。

(1) 非集計行動モデルのパラメータ推定方法

従来より、非集計行動モデルのパラメータ推定には最尤法が用いられている。それは、モデルを以下の式で表す場合、

$$P_{in} = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}} \quad (1)$$

$$V_{in} = \sum_{k=1}^K \theta_k X_{ink} \quad (2)$$

次の尤度関数を最大化するパラメータ θ を求める方法である。

$$L(\theta) = \sum_n \sum_i \delta_{in} \ln P_{in} \quad (3)$$

ただし、 i は選択肢、 n は個人、 P_{in} は選択肢 i を選ぶ確率、 V_{in} は選択肢 i の効用、 C_n は選択肢の集合、 X_{ink} は k 番目の説明変数、 θ_k はパラメータ、そして δ_{in} は個人 n が選択肢 i を選択していれば1、それ以外では0となる変数である。

以上で求められる最尤推定量 $\hat{\theta}$ は、以下の正規分布に漸近的に従うことが知られている。

$$MVN(\hat{\theta}, \Sigma) \quad (4)$$

$$\Sigma = (-\nabla^2 L(\hat{\theta}))^{-1} \quad (5)$$

(5) 式の分散共分散行列の各要素は、尤度関数の2階微分行列の逆行列にマイナス1を乗ずることで求められる。また(4)式は次のようにも記述できる。

$$P(\theta) = \frac{1}{(2\pi)^{\frac{K}{2}} |\Sigma|^{\frac{1}{2}}} \cdot \exp\left\{-\frac{1}{2} (\theta - \hat{\theta})^t \Sigma^{-1} (\theta - \hat{\theta})\right\} \quad (6)$$

(2) ベイズ推定によるパラメータ更新方法

本方法では(6)式で表されるパラメータ θ (以下では区別する必要から、これを θ^d と書く。)の分布を事前分布と考えた上で、更に追加情報が得られた状況を想定し、パラメータの更新を行うものである。追加的に得られる情報(集計データ)を Q^0 とおき、これに対して非集計行動モデルを集計して得られる値を $Q(\theta)$ とおく。

ここで、観測値 Q^0 が観測時の誤差や変動などを伴うために、誤差 ε を有すると考える。 ε の分布を平均ゼロの正規分布($MVN[0, \Sigma_0]$)と仮定すれば、 $Q^0 = Q(\theta) + \varepsilon$ において、 Q^0 の分布は、

$$P(Q^0 | \theta) = \frac{1}{(2\pi)^{\frac{J}{2}} |\Sigma_0|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} (Q^0 - Q(\theta))^t \Sigma_0^{-1} (Q^0 - Q(\theta))\right\} \quad (7)$$

と表される。 Q^0 ベクトルの次元 J は観測集計データの個数に対応し、誤差の分散共分散行列は $J \times J$ マトリックスである。

(7)式で与えられる Q^0 の分布は、パラメータ θ に関する尤度関数を意味する。

ベイズ推定では、新たにデータ Q^0 が追加された場合に、パラメータの事後分布 $P(\theta | Q^0)$ を、

$$P(\theta | Q^0) \propto P(\theta) \cdot P(Q^0 | \theta) \quad (8)$$

なる比例関係より求め、その最大化によってパラメータを決定する。ここでは、(6)、(7)式を用いて、

$$P(\theta | Q^0) \propto \exp\left\{-\frac{1}{2} (\theta - \theta^d)^t \Sigma^{-1} (\theta - \theta^d) - \frac{1}{2} (Q^0 - Q(\theta))^t \Sigma_0^{-1} (Q^0 - Q(\theta))\right\} \quad (9)$$

を導出でき、(9)式の最大化によって、新たなパラメータの推定値(最頻値)が求められる。

したがって、

$$\min_{\theta} [(\theta - \theta^d)^t \Sigma^{-1} (\theta - \theta^d) + (Q^0 - Q(\theta))^t \Sigma_0^{-1} (Q^0 - Q(\theta))] \quad (10)$$

を解けば良いことになる。この計算には先に非集計行動モデルのパラメータを推定した場合同様、逐次計算を要する。

(3) パラメータの近似解を求める方法

(10)式の最小化問題を直接解くことをしなくとも、求めるべきパラメータの近似解ならば容易に得ることが出来る。その方法を以下に記述する。

(9)式もしくは(10)式の第1項はパラメータ θ

に関して線形であるが、 $Q(\theta)$ が θ に関して非線形であるために第2項は非線形となっている。したがって、逐次計算による求解が必要となった。ここで、第2項も第1項同様に線形であれば(9)式自体が θ に関して正規分布に従うことを示すことが出来る。正規分布であれば平均値が最頻値と一致し、逐次計算によって最頻値を求める必要はなくなる。

そこで第2項の線形化を行うために、 $Q(\theta)$ を θ^d の回りで展開する。すなわち、

$$Q(\theta) \doteq Q(\theta^d) + \sum_{k=1}^K \frac{\partial Q(\theta^d)}{\partial \theta_k} (\theta_k - \theta_k^d) \quad (11)$$

と求められる。したがって、第2項の $Q^0 - Q(\theta)$ は、ベクトル表示によって次のように表せる。

$$Q^0 - Q(\theta) = F - G\theta \quad (12)$$

$$G_{jk} = \frac{Q_j(\theta^d)}{\theta_k} \quad (13)$$

$$F_j = Q_j^0 - Q_j(\theta^d) + \sum_{k=1}^K G_{jk} \theta_k^d \quad (14)$$

これを(9)式に代入すれば、

$$P(\theta | Q^0) \propto \exp\left\{-\frac{1}{2} (\theta - \theta^d)^t \Sigma^{-1} (\theta - \theta^d) - \frac{1}{2} (F - G\theta)^t \Sigma_0^{-1} (F - G\theta)\right\} \quad (15)$$

が得られる。さらに、 θ^r 及び Σ_r を以下のように定義すれば、

$$\theta^r = \theta^d + \Sigma G^t (\Sigma_0 + G \Sigma G^t)^{-1} (F - G \theta^d) \quad (16)$$

$$\Sigma_r = \Sigma - \Sigma G^t (\Sigma_0 + G \Sigma G^t)^{-1} G \Sigma \quad (17)$$

事後分布 $P(\theta | Q^0)$ は、

$$P(\theta | Q^0) \propto \exp\left\{-\frac{1}{2} (\theta - \theta^r)^t \Sigma_r^{-1} (\theta - \theta^r)\right\} \quad (18)$$

と書き改めることが出来、 θ^r が正規分布の期待値であり、 Σ_r が分散共分散行列を表すことが分かる。

以上より、(16)式を用いて θ^r を算出すれば、それが求めるべきパラメータの更新値になる。

4. 本方法の適用方法

3.では、非集計行動モデルを集計して得られる $Q(\theta)$ がどの様に与えられるかを述べていない。

本章では、この点を解説する。

(1) 各種集計方法への対応

集計方法については、数え上げ法、分類法、簡便法の何れであっても、本方法を適用可能である。与

えられた集計データが、モデルの選択肢毎の利用者数に一致する単純な場合を想定すると、数え上げ法であれば、サンプル数を N として、

$$Q_i(\theta) = \frac{T}{N} \sum_{n=1}^N P_{in} \quad (19)$$

また、分類法（簡便法も含む）では、 W_g を重みとして、

$$Q_i(\theta) = T \cdot \sum_g W_g P_{ig} \quad (20)$$

によって、各々 $Q(\theta)$ が定義される。ここで T は、

$$T = \sum_i Q_i^0 \quad (21)$$

によって与えられる総量のデータである。

したがって、(19) もしくは (20) 式を用いて、(9) 式の最大化を行えば、所要のパラメータを推定できる。

(2) 各種集計データへの対応

次に、どのような種類の集計データであれば、本方法を適用可能であるかを考える。 $Q(\theta)$ が非集計行動モデルの集計値であるため、 Q^0 がモデルから集計可能な量で有りさえすれば、理論上は適用可能である。このことを具体的に以下に解説する。

a) 母集団全体の集計データが既知の場合

(1) に述べたように、モデルの選択肢別に母集団における集計量が得られていれば、何ら工夫を要せずに推定が行える。また、選択肢集合の総てではなく、一部（2 つ以上の選択肢）の集計データが得られている場合でも、ロジットモデルの IIA 特性を利用してモデルパラメータを更新できる。例えば、 a 、 b 2 つの選択肢の集計データ（利用者数）が得られているとき、 a 、 b 間の確率 $P_n(i|a, b)$ は、

$$P_n(i|a, b) = \frac{e^{V_i}}{e^{V_a} + e^{V_b}}, \quad i=a \text{ or } b \quad (22)$$

で表される。したがって、集計量 Q_a^0 、 Q_b^0 に対応するモデルからの推計値 $Q_i(\theta)$ が、分類法では、

$$Q_i(\theta) = (Q_a^0 + Q_b^0) \sum_g W_g P_{ig}(i|a, b) \quad (23)$$

と書けるので、これよりパラメータを更新できる。また同様に考えれば、選択肢を幾つも寄せ集めた集計量が既知である場合等にも同様な方法で適用できる。小ゾーン単位で作成した目的地選択モデルにおいて、大ゾーンの集中交通量が既知である場合等が本例に相当する。

b) セグメント集団の集計データが既知の場合

交通機関選択モデルの構築において、特定 O D 間の手段別交通量が得られる場合や、目的地選択モデルにおいて、特定 O D 間の分布交通量が既知である場合など、対象地域の中の部分的なデータが利用可能なケースでも、それらの推計値をモデルから集計して求められさえすれば、本方法を適用することが出来る。

(3) 本方法の適用対象

本方法を生かせる適用対象として、非集計データに加えて、O D 間の手段別交通量、O D 交通量、集中交通量の一部もしくは総てが得られている場合に、交通手段選択、目的地選択モデルのパラメータを推定する例を代表として挙げられることは、以上の説明から明らかである。このとき、集計量データには、国調、交通量観測値、駅乗降客数、入り込み者数、P T 調査等のデータを採用すれば良い。

具体的には、入り込み者数と個人データもしくは非集計モデルパラメータとの統合利用や、国調データと非集計データとの統合利用、都市間の交通機関別旅客数と非集計データとの統合利用などに活用することが考えられる。

5. 本推定方法の特性

3. で開発したパラメータ更新方法の有効性を確認する目的で、本章ではシミュレーション分析を行った。モデルパラメータの更新に影響を及ぼす要因は、1) 非集計モデルパラメータの信頼性、2) 集計データの種類、3) 集計データの信頼性の 3 つに代表されるが、ここでは特に 2)、3) に着目して分析している。

(1) 使用データの概要

本章のシミュレーション分析に用いたデータは、昭和 52 年に前橋高崎都市圏で実施された P T 調査データである。P T データの内、前橋市内にトリップの発着地がある通勤目的で、かつバス・自動車・二輪車・徒歩の 4 手段を利用したサンプル 4566 票を抽出して分析対象とした。このサンプルにおける各手段利用者数は、バス 333、自動車 2422、二輪車 1213、徒歩 578 であり、これを 3. に示した事後情報 Q^0 の値と仮定した。更に、そこから再抽出した 484 サンプル（バス 32、自動車 245、二輪車 139、徒歩 68）を非集計行動モデルの作成用データとして、

モデルパラメータの推定に用いた。

(2) シミュレーション分析の条件設定

次に分析の前提となる各種条件を以下に説明する。

まず、事後情報の精度を表す分散共分散行列には、その大小で結果がどの様になるかを調べるために、次に表される数値を設定した。 V_{ij} を行列要素として、

$$V_{ij} = Q_i^0 \times Q_j^0 \times \alpha \quad (i=j) \\ = 0 \quad (i \neq j) \quad (24)$$

ここで、 α は信頼性の程度を表す指標となり、事後情報の変動係数の2乗に一致する。

また、非集計モデルのパラメータ θ^d をそのまま用いて推計される集計値 $Q(\theta^d)$ と観測値 Q^0 との差が異なる幾つかのケースに対して分析を行う必要から、 Q^0 として(1)に示した値の他に、次に示す3種類のデータを設定した。(表-1) 表中に示した $\Delta Q(\theta^d)$ は、(26)式の θ^r の替わりに θ^d を用いて計算した集計誤差(これを初期集計誤差と称する)を表す。これらは現実に本方法を適用する際に、上記両者が様々に異なる可能性を考慮したものである。

パラメータの更新結果を把握するための指標として、パラメータの変化率を以下の式で、

$$\Delta \theta = \sum_{k=1}^K \left| \frac{\theta_k^d - \theta_k^r}{\theta_k^d} \right| \times \frac{100}{K} \quad (\%) \quad (25)$$

更に、集計量の差を以下の式で定義した。

$$\Delta Q(\theta^r) = \sum_{j=1}^J \left| \frac{Q_j^0 - Q_j(\theta^r)}{Q_j^0} \right| \times \frac{100}{J} \quad (\%) \quad (26)$$

ここで、 K はパラメータ数を J は事後情報として与えられる集計データの数を表し、 θ^r 及び $Q(\theta^r)$ は更新されたパラメータとそれを用いて推計される集計値を各々表している。従って、(25)式は更新前の非集計行動モデルのパラメータと更新後のパラメータとの相対的な変化の割合を示し、(26)式は観測された集計量とパラメータ更新後の集計量との相対的な差(これを事後集計誤差と称する)を表している。

(3) パラメータ更新値の基本的挙動

ここでは、確認のために本方法によるパラメータ更新値がどのような挙動を示すかを計算した。まず、表-2に484サンプルで推定した交通手段選択モデルのパラメータ一覧を示す。

表-1 各ケースの手段別利用者数

ケース 選択手段	ケース0	ケース1	ケース2	ケース3
バス	333	457	685	913
自動車	2442	2283	2055	1826
二輪車	1213	1141	913	685
徒歩	578	685	913	1142
$\Delta Q(\theta^d)$	8.44%	14.08%	35.44%	57.18%

表-2 非集計交通手段選択モデルのパラメータ推定結果(バス、自動車、二輪車、徒歩)

説明変数	パラメータ	t 値
1 所要時間(分)	-0.1162	25.82
2 自動車保有台数 (千人/人)	0.001348	2.97
3 免許保有ダミー	3.191	14.09
4 二輪車保有台数 (千台/人)	0.001917	6.20
5 バス定数項	-1.534	9.22
6 自家用車定数項	-3.534	15.42
7 二輪車定数項	-1.592	6.86
尤度比	0.339	サンプル数 484

表-3 モデルパラメータの更新結果(ケース1)

パラメータ	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
1.	-0.1160 (-0.12)	-0.1160 (-0.12)	-0.1155 (-0.68)	-0.1155 (-0.68)	-0.1155 (-0.68)	-0.1154 (-0.71)
2.	0.001350 (0.12)	0.001343 (-0.40)	0.001321 (-2.09)	0.001386 (2.75)	0.001385 (2.70)	0.001469 (8.89)
3.	3.186 (-0.14)	3.186 (-0.14)	3.166 (-0.77)	3.166 (-0.77)	3.166 (-0.77)	3.166 (-0.77)
4.	0.001927 (0.51)	0.001899 (-0.91)	0.001831 (-4.46)	0.001947 (1.55)	0.001918 (0.06)	0.001897 (-1.01)
5.	-1.498 (-2.37)	-1.498 (-2.37)	-1.316 (-14.20)	-1.072 (-30.12)	-1.072 (-30.12)	-1.072 (-30.12)
6.	-3.536 (0.06)	-3.536 (0.06)	-3.556 (0.65)	-3.598 (1.82)	-3.598 (1.82)	-3.598 (1.82)
7.	-1.611 (1.20)	-1.611 (1.20)	-1.695 (6.49)	-1.781 (11.90)	-1.781 (11.90)	-1.781 (11.90)

上段：パラメータ

下段：変化率(%)

この結果を基にして、3.(2)で提案した逐次計算によって求めたパラメータの更新結果を、ケース1について表-3にまとめた。表中の下段はパラメータの変化率を示したものである。 α が小さい所でパラメータがより大きく修正されている様子が伺える。

また、表-4はこの場合の集計結果を表にまとめたものである。これを見ても、集計データの信頼性が高い場合に事後集計誤差が殆ど無くなることが分かる。

以上を他のケースも含めて、(25)式、(26)式の平均変化率で図示したものが図-1、2である。横軸はログスケールで表した α の値であり、図-1の縦軸はパラメータの変化率を、また図-2では集計値の変化率を各々表している。 α が小さい、即ち事後情報の信頼性が高い場合に、全般的にパラメータが大きく修正され、また集計量が与えた数値に近付いていることが分かる。また、更新前に集計誤差が大きなケースでは、パラメータの変化がより大きいことが読み取れる。信頼性の低い場合にパラメータが殆ど修正されない点も明らかであり、本方法の基本的な挙動が確認できたと言える。

(4) 近似解法の有効性の検討

3.(3)に提案した、パラメータの近似解を逐次計算なく求める方法(近似解法と称する)は、パラメータに関する線形化のために、一次近似を用いている。従って、パラメータが大きく修正される場合に、パラメータの推定誤差が無視できなくなる可能性がある。

本節では、この点を(3)と同様な方法で検討した。図-3、4は近似解法によって推計した結果を図示したものである。図-3に示したパラメータの変化傾向は先の図-1と同様であることが分かる。両者の違いはケース1、2の初期集計誤差が少ない場合には殆ど無い。一方、図-4に表した事後集計誤差の値は、 α が 10^{-4} 以下で再び大きくなっており、その傾向は初期集計誤差の大きいほど著しい。

この様な傾向の生じた理由は、明らかに線形近似にあり、パラメータの変化が激しいとき、近似解法を使う場合には注意を要することがわかる。ただし、 α が 10^{-4} とは、変動係数の1%に相当するが、その様な高い精度のデータは一般には存在しないと考えられ

表-4 手段別集計値と事後集計誤差(ケース1)

パラメータ	α	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
バス		3123	3142	3797	458.4	460.7	458.9
自動車		23082	23127	23047	2263.4	2268.7	2286.7
二輪車		13008	12894	11974	1168.1	1155.8	1139.9
徒歩		6446	6496	684.1	675.9	680.7	680.4
$\Delta Q(\theta^r)$		13.1%	12.7%	5.7%	1.2%	0.9%	0.4%

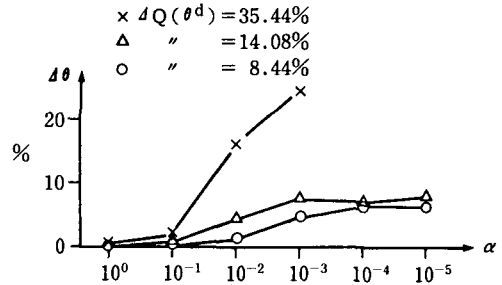


図-1 パラメータの変化と集計データの特性

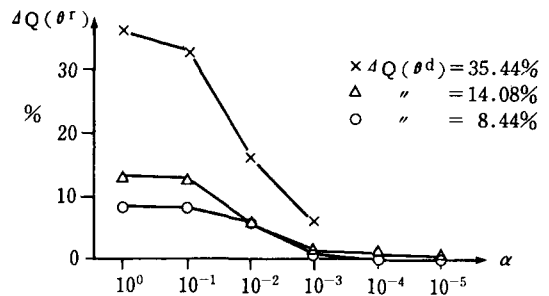


図-2 事後集計誤差の変化と集計データの特性

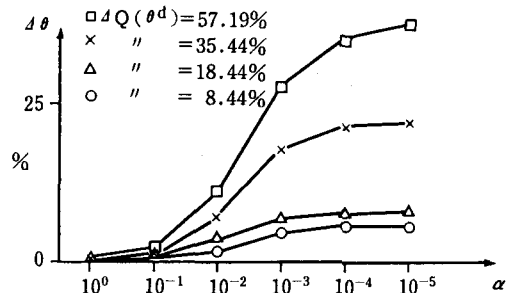


図-3 パラメータの変化と集計データの特性
(近似解法の場合)

る。

また、パラメータを大きく変化させる要因は、初期集計誤差と事後情報の信頼性との2つに代表されるが、これらは両者とも既知であるため、その大きさ次第で逐次解法に移行する判断をすれば良いと考える。

(5) 一部の集計データが既知の場合の検討

先の例で、一部分のデータが収集されている場合のパラメータ更新について、簡単な確認作業を行った。バスの利用者数と全体の需要量との2つのデータが既知として、(4)で確認した近似解法によって推計した集計量を表-5に示す。バス利用者数の事後集計誤差は明らかに改善され、観測値である333サンプルに近付くが、他の交通手段については変化が明確には現れていないことが分かる。

(6) 他のパラメータ推定方法との比較

本節では、他の異なる方法と比較することによって、本方法の特徴をより明確にすることを考えた。ここでいう他の方法とは、森地・屋井(1985)の方法と選択肢別標本抽出法に対して開発されたパラメータ推定方法との2つである。

a) 定数項修正方法としての比較

本研究で開発した方法は、パラメータの一部の更新にも用いることが出来る。コンスタント項を除く変数のパラメータを定数と考えて、修正計算に取り込まないことにすれば、本方法は集計データを用いてコンスタント項を修正する方法と言える。森地・屋井(1985)は既に、集計データを用いてコンスタント項を修正する方法を提案しているが、それと本方法とがどのような関係にあるかを調べた。表-6は α を 10^{-5} と置いた上で、逐次計算によって更新したコンスタント項を示したものである。コンスタント項以外は定数として扱い、修正を施していない。また、同表の下段は森地・屋井によって求めた値である。これを見れば両者が殆ど一致していることがわかる。本研究の方法は集計データの分散がゼロである場合には適用できないが、逆に森地・屋井の方法では分散がゼロの確定値に対してのみ適用可能である。したがって、森地・屋井(1985)は、本方法で言うところの事後情報の精度が限りなく向上した極限での解を表していると考えることが出来る。

b) 選択肢別標本抽出法との関連

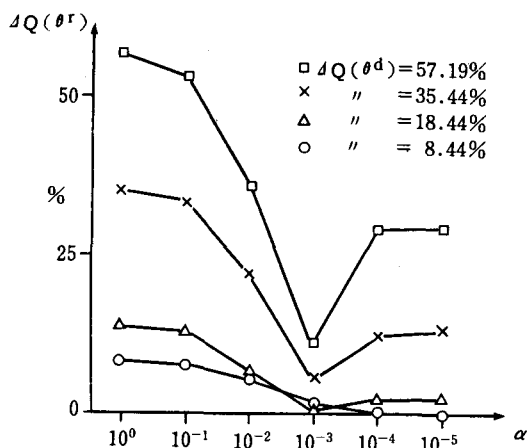


図-4 事後集計誤差の変化と集計データの特性 (近似解法の場合)

表-5 利用者総数とバス利用者数が既知の場合の集計結果(ケース0)

選択肢 \ α	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
バス	3022	3054	319.9	334.2	332.0	334.5
自動車	23120	23090	22959	22832	22852	2283.2
二輪車	13103	13087	13014	12944	12954	1293.7
徒歩	6414	6428	6486	6541	6532	654.5

表-6 定数項修正方法としての比較結果

ケース \ 選択肢	ケース1	ケース2
バス	-1.094 -1.093	-0.952 -0.952
自動車	-3.608 -3.607	-4.190 -4.190
二輪車	-1.806 -1.805	-2.374 -2.373

- ・上段は本方法($\alpha = 10^{-5}$)、下段は森地・屋井(1985)。
- ・徒歩のコンスタント項はゼロに固定。

選択肢別標本抽出法では、選択肢別にサンプルを収集したためにサンプル内の分担率が実際と異なっている場合のパラメータ推定問題を扱っている。実際の分担率の値を追加情報として用いることで統計的に一致性のある推定量と出来ることが証明されている。

分担率の比を重みとして用いて尤度関数を構成する、WESML (Weighted Exogenous Sampling Maximum Likelihood) 推定²⁾を例に取ると、非集計行動モデルのパラメータ推定時に集計データを用いる点と、分担率の信頼性を考慮することができない点とで、本方法と大きな違いがあると言える。

表-7に初期集計誤差が14.08%の場合のWESML推定量の値を示した。同表の右側2列には、参考として α が 10^{-5} と 10^0 の2つについて本方法で更新したパラメータの値を記してある。これより、両者のパラメータにはさほど大きな差が無いことが読み取れる。ただし、WESML推定では一般に、集計データを総ての選択肢について収集する必要がある、また使用できるデータの種類も選択肢毎の全体シェアに限られるため、本来の目的以外に応用する可能性は少ない。その点でも、本研究の方法は応用性の高いものと考えられる。

表-7 選択肢別標本抽出法におけるパラメータ推定との比較結果(ケース1)

推定方法 パラメータ	WESML推定	本方法 ($\alpha=10^{-5}$)	本方法 ($\alpha=10^0$)
1.所要時間	-0.1161	-0.1154	-0.1160
2.自動車保有	0.001240	0.001469	0.001350
3.免許保有	3.298	3.166	3.186
4.二輪車保有	0.001946	0.001897	0.001927
5.バス定数項	-1.197	-1.072	-1.498
6.自動車定数項	-3.676	-3.598	-3.536
7.二輪車定数項	-1.831	-1.781	-1.611
$\Delta\theta$	-	5.84%	7.76%

・ $\Delta\theta$ はWESMLとの相対差((25)式)

6. おわりに

本研究は非集計行動モデルのパラメータを推定する新たな方法を提案したものであり、研究の成果を

まとめれば、以下ようになる。

(1) 集計データと個人データとの両者を用いて非集計行動モデルのパラメータを推定する一般性の高い方法を開発し、その有効性をシミュレーション分析を通して確認した。

(2) 基本的な方法に加え、推定の容易な近似解法を開発し、その特性と適用上の限界を明らかにした。

(3) 既存の方法と本研究の方法とを比較して両者の関連を明らかにした。

なお、本研究で提案した方法は、モデル構築段階のパラメータ修正に留まらず、PT調査の中間年次におけるモデルの更新や、他の地域へのモデルの移転時や、更に対象地域の一部のデータで作成したモデルを全域に拡大して利用する目的等に対しても適用可能と考えられ、移転可能性の成立条件の議論に新たな展開が期待される。また、現実の様々な状況に対して本方法を適用し、その有効性をより明確にして行く必要性も高く、その点は今後の課題である。

参考文献

- 1) 森地 茂, 屋井鉄雄, 田村 亨: 非集計行動モデルによるOD交通量推計方法, 土木計画学研究・論文集, No. 2, pp. 45-52, 1985. 1.
- 2) 森地 茂, 屋井鉄雄: 非日常的交通への非集計行動モデルと選択肢別標本抽出法の適用性, 土木学会論文報告集, No. 434, pp. 161-170, 1984. 3.
- 3) 森地 茂, 屋井鉄雄, 田村 亨: 非集計交通手段選択モデルの地域間移転可能性, 土木学会論文集, No. 359, pp. 107-115, 1985. 7.
- 4) 森地 茂, 屋井鉄雄, 平井節生: 非集計行動モデルのパラメータ推定における集計データの利用方法, 第42回土木学会年次学術講演会講演概要集4, 1987. 9.
- 5) Morichi, S. and T. Yai: Prediction of Trip Distribution by Disaggregate Behavioral Model, Proceedings of the World Conference on Transport Research, Vol. 2, pp. 1582-1599, 1986. 5
- 6) Quevedo, S. L. G.: Combining Survey and Aggregate Data for Model Estimation, 1985. 6.
- 7) Ben-Akiva, M.: Methods to Combine Different Data Sources and Estimate Origin-Destination Matrices, Draft Paper, 1987.