土木学会論文集 No. 470／IV-20, pp. 125~133, 1993. 7

特集論文(交通行動分析の新展開)

複数データの統合による交通需要モデルの推定技法

屋井鉄雄*・森地　茂**・魚谷　憲***

本研究では，複数の異なるデータが存在する場合に，それらを組み合わせて交通需要モデルを作成する方法論を開発し，その特性を論じた．異なるデータの統合利用は限られた情報の有効利用という利点から，従来より幾つか検討されてきた．本研究ではこれらを体系的に整理するとともに，非集計モデルのパラメータ更新時の集計データ活用に着目して新しい方法論を示している．休日道路交通に対する適用を通して各方法の特性を概略把握した．

Key Words : *bayesian updating, discrete choice model, aggregate data, dirichlet distribution, route choice behavior*

1. はじめに

複数の異なるデータソースを同時に用いて交通需要予測を行う方法論は，需要分析の活用領域を増せる利点のため，従来より様々な場面で用いられてきた．たとえば，1970年代後半より実用化が進められた非集計モデルでは，調査，推定，移転の各予測段階に複数のデータを用いる方法が提案されている．これらは，より精度の高い予測の実現を目的として提案され，既に実用的にも活用されてきた．近年，交通計画の対象が広がりを見せるにつれ，データ蓄積が十分でない様々な状況が顕在化し，この様な技術も踏まえ，新たな交通需要予測法を確立する必要性も強まってきている．比較的データ整備が遅れた休日交通や都市間交通，データ整備環境が劣る国際交通や発展途上国の交通，現存しない新しい交通サービス，また，従来より指摘される大規模調査の希な地方中小都市の交通などである．

本論文では，非集計分析に関わる従来のデータ統合技術を最近の展開も含めて体系的に整理した上で，筆者らが進めてきた方法論を新たに発展させる．そして，その特性と適用対象とを示し，簡単なケーススタディによって実用可能性を検討する．したがって，本研究の目的は従来の技法の整理と新たな方法論の提案との２つである．新しく提案する方法は，少量の調査サンプルで作成した離散選択非集計モデルを，ベイズ更新の考え方に従って交通量など追加的に得られる集計データで更新する方法である．これにより，現況再現性と予測精度の向上が期待でき，離散選択モデルの交通需要分析における適用範囲が拡大できる．集計データに現れた集団としての交通現象と個々のトリップの情報とを同時に用いることで，データ不足で表現が困難な現象の記述や，交通現象の理解を助ける方法として活用できると考えられる．アンケート調査と交通量カウント調査や既存交通統計との組合わせ調査法や，それらのデータを用いた適切な推定法，またモデルの他地域への移転・拡大利用法など，予測作業上，有効な方法論確立が期待されている．本論文では，集計量の分布を従来提案したものと変えることによって，より広範な適用可能性を探ることに特色がある．

2. 従来のデータ統合方法の体系的整理

(1) データ統合手法の発展の経緯

複数の種類のデータを用いて需要予測を行う方法論は，非集計モデルの開発当初より散見できる．Athertonら（1976）は，非集計モデルの移転方法として，移転する元のモデルのパラメータのうち，スケールと定数項との２種を，移転される地域の非集計データまたはゾーン間集計シェアを用いて再推定する方法を示した．

以下の式で x_i は説明要因，C は選択肢集合，P_{in} は個人 n が選択肢 i を選ぶ確率であり，α, β が新しいパラメータである．

$$P_{in}=\frac{\exp(\alpha\theta x_i+\beta_i)}{\sum_{l\in C}\exp(\alpha\theta x_l+\beta_l)} \quad \cdots\cdots(1)$$

また，２つの地域それぞれでモデル作成に足る非集計データが得られていれば，各々でモデルパラメータを推定した後，ベイズ更新法により新たなパラメータを推定する方法も提案された．これはパラメータが漸近的に正規分布する性質を利用した方法であり，分散の大小で重み付けした更新法になっていることがわかる．

$$\theta'=\Sigma'(\Sigma_0^{-1}\theta_0+\Sigma_1^{-1}\theta_1) \quad \cdots\cdots(2)$$

$$\Sigma'=(\Sigma_0^{-1}+\Sigma_1^{-1})^{-1} \quad \cdots\cdots(3)$$

* 正会員　工博　東京工業大学助教授　工学部土木工学科
（〒152　目黒区大岡山 2-12-1）
** 正会員　工博　東京工業大学教授　工学部土木工学科
*** 正会員　運輸省（元東京工業大学修士課程学生）

上の式で θ_0, Σ_0 は移転する元のモデルのパラメータとその分散共分散を，θ_1, Σ_1 は移転される側の地域の少数のデータより推定したモデルのパラメータとその分散共分散を，θ', Σ' は更新された新たなパラメータとその分散共分散とを各々示す．

また，非集計分析では複数の異なる調査を行ってモデルを作成する方法が，選択肢別標本抽出法として確立されている．たとえば，Lerman & Manski（1979）らによる，以下の重み付き尤度関数WLの最大化によって，非集計モデルのパラメータを推定する方法がある．これも複数の選択肢別の抽出サンプルと各選択肢の集計シェアとを同時に用いる統合推定方法の１つである．

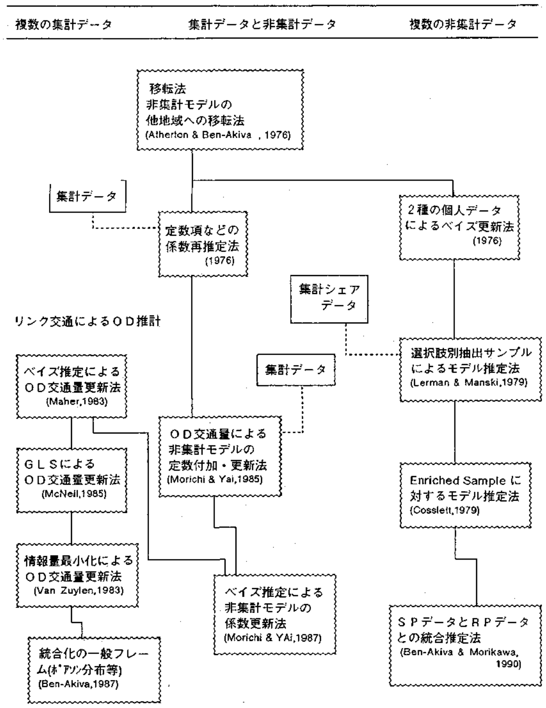
$$WL=\sum_{i=1}^{C}\sum_{n=1}^{Ni}\frac{Q(i)}{H(i)}\cdot\ln P_{in} \quad\cdots\cdots(4)$$

ここで，N_i は選択肢 i ごとのサンプル数，$Q(i)$ は選択肢 i の実績シェア，$H(i)$ は推定に用いたサンプルにおける選択肢 i のシェアを表している．

また，家庭訪問調査のサンプルと選択肢別調査のサンプルとを同時に収集してモデル作成を行う，いわゆる enriched sampling の方法論も示されている．それによれば両者のサンプルを同時に用いたプーリング推定が行われる．

1980年代に入って，データ統合を必要とする機会は，大規模調査の減少，新しい予測対象の出現などによって増してきた．たとえば，リンク交通量からOD表を推計する方法としても，いくつもの検討が進められている．Maher（1983）はベイズの定理を用いて事後的に得られたリンク交通量によりOD交通量を更新する方法を示した．また，McNeil ら（1985）は同様な問題を一般化最小２乗法によって定式化している．その他，情報量の最小化によってOD交通量を修正する方法を示した Van Zuylen ら（1980）や，OD交通量と配分交通量との同時推定問題を扱った Fisk ら（1983）の研究がこの時期に集中している．

一方，1980年代半ばになって，いったん実用化を終えた非集計モデルに対して，再び方法論を発展させる試みが始められた．そのうち，複数のデータソースを同時に扱う方法論は，非集計分析の利点を生かしながら適用範囲を広げる工夫として大いに期待された．Quevedo（1985）では，連続型の選択モデルを正規分布関数として定義した上で，追加的に得られる集計データとの同時尤度関数の最大化によって，モデルパラメータを推定する方法を示した．また，OD交通量の推計に対して，森地ら（1985）は非集計モデルに定数項を加えて推計する方法を示し，Ben-Akiva（1987）は調査データと交通量観測データとの両者を，ポアソン分布に従う同時尤度関数の最大化によって統合し，OD交通量の推計を行う方法論を示している．



図—1　複数データの統合利用法の発展経緯

さらにこの時期，森地ら（1987）はベイズの定理を応用して，追加する集計データが正規分布に従うと仮定した上で，非集計モデルのパラメータを更新推定する方法論を示した．一方，Ben-Akiva ら（1990）は選択実績データと選好データという性質の異なる複数のデータソースを効用関数の誤差項の相違によって表し，両データよりパラメータ推定を行う方法を示した．

以上の方法論開発の経緯を図—1のようになる．統合するデータ種別で３つに大別すれば，①複数の集計データ，②集計データと非集計データ，③複数の非集計データに分かれる．このうち，筆者ら（1987）の方法では，追加的に得られる集計データを加えることで，信頼性や再現性のより高いモデルを得られると考え，既に地下鉄延伸の免許申請のための需要予測などで実用的に用いられてきた[5]．以下では本論文の方法論開発にあたり出発点となるこの更新方法を概観する．

（２） ベイズ更新法による複数データの統合

本方法では，１度作成したモデルのパラメータを任意の時点や場所における集計データで更新できる点と，計算上の安定性とを重視して，段階推定法であるベイズ更新法を用いている．

離散選択モデルのパラメータ θ^d が，個人データだけを用いて最尤推定された場合を考える．たとえば，ロジットモデルを考えると，

$$P_{in}=\frac{\exp(V_{in})}{\sum_{l\in C}\exp(V_{ln})}\quad\cdots\cdots(5)$$

であり，ここで効用関数 V_{in} が通常以下の様にパラメータを伴う説明要因の線形関数で表される．

$$V_{in}=\sum_{k=1}^{K}\theta_k^d\cdot X_{ink} \quad (6)$$

このパラメータベクトル θ^d は漸近的に正規分布に従うため，以下のような密度関数を持つ確率変数と見なすことができる．

$$f(\theta)\propto\exp\{-(\theta-\theta^d)^t\Sigma_d^{-1}(\theta-\theta^d)/2\} \quad (7)$$

これは平均 θ^d で分散共分散マトリクスが Σ_d の多項正規分布である．これをパラメータの事前分布と想定することによって，ベイズ更新の手続きにより新たな情報の追加された条件でパラメータの更新を行える．

ここで，集計データ行列が Q で与えられ，その値の離散選択モデルを用いた推計値を $Q(\theta)$ で表し，Q を正規分布に従う誤差 ε を用いて，

$$Q=Q(\theta)+\varepsilon \quad (8)$$

と表せば，追加情報である尤度関数は，

$$L(Q|\theta)\propto \exp\{-(Q-Q(\theta))^t\Sigma_q^{-1}(Q-Q(Q))/2\} \quad (9)$$

によって表される．ここで，Σ_q は誤差の分散共分散行列である．

個人データのみから最尤法で推定したパラメータ θ の分布を事前分布 $f(\theta)$ と考えると，事後確率密度 $f(\theta|Q)$ は，

$$f(\theta|Q)\propto f(\theta)\cdot L(Q|\theta) \quad (10)$$

によって表される．これに式（7），式（9）を代入すれば，結局，

$$f(\theta|Q)\propto \exp\{-(\theta-\theta^d)^t\Sigma_d^{-1}(\theta-\theta^d)/2 -(Q-Q(\theta))^t\Sigma_q^{-1}(Q-Q(\theta))/2\} \quad (11)$$

によって事後確率密度が表される．この式の最大化，つまり，カッコ内の絶対値の最小化によって，新たなパラメータベクトル θ を決定できる．

$$\min_\theta\{(\theta-\theta^d)^t\Sigma_d^{-1}(\theta-\theta^d) +(Q-Q(\theta))^t\Sigma_q^{-1}(Q-Q(\theta))\} \quad (12)$$

$Q(\theta)$ は，集計法に簡便法を用いるなら，

$$Q_i(\theta)=T\cdot P(i|X,\theta),\ i=1,\cdots,m \quad (13)$$

である．ここで，T はマーケットボリューム，P は説明要因の平均値ベクトル X を用いた選択肢 i のシェア，m は選択肢数である． Qは離散選択モデルを用いて推計できる値であれば良い．

また，パラメータの変化が小さいことを前提とすれば，$Q(\theta)$ を θ^d のまわりで線形近似することによって，簡便にパラメータ更新を行うことができる．すなわち，

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

$$G_{jk}=\frac{\partial Q_j(\theta^d)}{\partial\theta_k} \quad (15)$$

ここで，

$$F_j=Q_j-Q_j(\theta^d)+\sum_{k=1}^{K}G_{jk}\theta_k^d \quad (16)$$

とおけば，式（11）は結局，以下のように表される．

$$f(\theta|Q)\propto\exp\{-(\theta-\theta^r)^t\Sigma_r^{-1}(\theta-\theta^r)/2\} \quad (17)$$

$$\theta^r=\theta^d+\Sigma_dG^t(\Sigma_q+G\Sigma_dG^t)^{-1}(F-G\theta^d) \quad (18)$$

$$\Sigma_r=\Sigma_d-\Sigma_dG^t(\Sigma_q+G\Sigma_dG^t)^{-1}G\Sigma_d \quad (19)$$

上の式は期待値 θ^r を持つ正規分布であるため，この値 θ^r をパラメータの点推定値とすれば良い．

なお，追加データの持つ分散 Σ_q の与え方については，その後，大規模調査のサンプリング誤差より求めたものや，データの変動を基に設定したものなどが提案されている（Yai et. al. (1989)）．

3. 本研究で提案する新しい統合方法

以上の更新方法では，追加される集計データの分布を多項正規分布とし，データの信頼性や選択肢間の競合関係をデータの分散共分散項で表す形式を採用した．しかし，集計データの誤差の大きさが不明な場合や，小さな交通量を扱う場合など，的確な分散設定が困難な場合にもデータ統合を行うための別の工夫が出来れば有意義であろう．需要モデルにシェアモデルを用いていることから，選択肢間の関係を直接考慮した分布の設定が必要になる場合も生じると考えられる．以上を考慮して本研究では次に示す2種類の更新法を新たに提案した．

（1） ディリクレ分布を用いた更新方法の提案

多項分布とよく似た分布にディリクレ分布がある．乗数部分がパラメータとして扱われる同時確率関数である多項分布に対して，ディリクレ分布ではシェアがパラメータとなる確率密度が考えられている．

交通量 $x_1,\cdots,x_i,\cdots,x_m$ が，パラメータ $(q_1,\ \alpha),\cdots,(q_i,\ \alpha)\cdots,(q_m,\ \alpha)$ の独立なガンマ分布，

$$f(x_i,q_i,\alpha)=\frac{1}{\alpha\Gamma(q_i)}\left(\frac{x_i}{\alpha}\right)^{q_i-1}\exp(-x_i/\alpha),\ x_i>0 \quad (20)$$

に従うと仮定する．これは交通量が正の数のみを取ることから，少ない交通量に対してはより現実的な仮定になる．このとき，その比 z_i,

$$z_i=x_i/(x_1+\cdots+x_m),\ i=1,\cdots,m-1 \quad (21)$$

の同時確率密度関数は，

$$f(z,q)=\frac{\Gamma(q_1+q_2+q_3+\cdots+q_m)}{\Gamma(q_1)\cdots\Gamma(q_m)}\cdot z_1^{q_1-1}\cdots z_{m-1}^{q_{m-1}-1}(1-z_1-\cdots-z_{m-1})^{q_m-1} \quad (22)$$

なるディリクレ分布に従うことが知られている．

ここで，x_i の期待値と分散は，

$$E(x_i)=q_i/\alpha \qquad (23)$$

$$E(x_i^2)=q_i/\alpha^2 \qquad (24)$$

であり，z_i の期待値，分散共分散は，

$$E(z_i)=q_i/\sum_{j=1}^{m}q_j \qquad (25)$$

$$E(z_i^2)=q_i(q_i+1)/\left\{\sum_{j=1}^{m}q_j\left(1+\sum_{j=1}^{m}q_j\right)\right\} \qquad (26)$$

$$E(z_iz_j)=q_iq_j/\left\{\sum_{j=1}^{m}q_j\left(1+\sum_{j=1}^{m}q_j\right)\right\} \qquad (27)$$

である．したがって，α を1.0に先決すれば q_i が交通量 x_i の期待値を表すことがわかる．ここで q_i が，選択モデルの真のパラメータを持つ以下の式で表されると仮定する．

$$q_i=T\cdot P(i|\theta) \qquad (28)$$

この式で，T は総交通量（$T=\sum x_i$），$P(i|\theta)$ は，式（5）のロジットモデルで表されると考える．

ここで，先の式（9）を用いて非集計モデルのパラメータを更新することを考える．式（9）と同様にして，

$$f(\theta|Q)\propto f(\theta)\cdot f(z,q) \qquad (29)$$

と表す．この両辺を対数変換し，式（7）と式（22）を代入すれば，

$$\ln f(\theta|Q)\propto -\frac{1}{2}(\theta-\theta^d)^t\Sigma_d^{-1}(\theta-\theta^d)$$

$$+\sum_{i=1}^{m}(q_i-1)\ln z_i-\sum_{i=1}^{m}\Gamma(q_i) \qquad (30)$$

を得る．この式の最大化によって新しいパラメータを得ることが出来る．本更新法の特徴は，交通機関分担等，選択肢間の相関関係をシェアの分布として明示的に用いた点と，集計データの分散をデータ別に確定したところにある．分散固定は強い仮定であるが，集計データが小さい場合にはその分散構造が未知であることも多く，通常用いられるポアソン分布と同様，簡便な活用が行い易い方法と言える．

また，複数の組の集計データ（$n=1,\cdots,N$）が得られている場合には，式（30）は，

$$\ln f(\theta|Q)\propto -\frac{1}{2}(\theta-\theta^d)^t\Sigma_d^{-1}(\theta-\theta^d)$$

$$+\sum_{n=1}^{N}\left\{\sum_{i=1}^{m}(q_{in}-1)\ln z_{in}-\sum_{i=1}^{m}\Gamma(q_{in})\right\} \qquad (31)$$

と書き換えられる．これはより一般性を有した推定式となる．このように，複数の観測交通量データが得られた場合を想定し，パラメータ分布の最頻値を求めれば，事後分布の形状が極度に歪んでいない限り，その値をもってモデルパラメータの更新値とできる．

また，q_i が交通量であることから整数分布を当てはめることも考えられる．このとき，$\Gamma(q_i)=(q_i-1)!$ が成り立ち，ガンマ分布はポアソン分布と等しくなる．すなわち，交通量の分布は期待値，分散が共に q_i のポアソン分布となる．この関係を式（30）に直接代入すれば，

$$\ln f(\theta|Q)\propto -\frac{1}{2}(\theta-\theta^d)^t\Sigma_d^{-1}(\theta-\theta^d)$$

$$+\sum_{i=1}^{m}(q_i-1)\ln z_i-\sum_{i=1}^{m}\sum_{k=1}^{q_i}\ln(q_i-k) \qquad (32)$$

を得る．この式の最大化によって新たなパラメータを求めることができるが，この場合，q_i を整数として計算を進める必要がある．

（2） 構造化された多項分布を用いた更新法

交通量 $x_1, x_2, x_3, \cdots, x_m$ が，以下の多項分布に従って得られると考える．

$$f(x,s)=\frac{x!}{x_1!\cdot x_2!\cdots x_m!}\cdot$$

$$s_1^{x_1}\cdots s_{m-1}^{x_{m-1}}\cdot(1-s_1-\cdots-s_{m-1})^{x_m} \qquad (33)$$

ここで，x は総交通量（$\sum x_i$），s には未知なパラメータ θ を含むシェアの関数を考える．この仮定によっても，同様なモデルパラメータの更新方法を作り出せる．上式が追加データによる尤度関数に相当するため，複数のODペア N で交通手段 i の交通量 $x_{in}(i=1,\cdots,m;n=1,\cdots,N)$ が観測された場合を想定すると，尤度 $L(x|\theta)$ は，

$$L(x|\theta)=\prod_n f(x_n,s_n) \qquad (34)$$

である．したがって事後確率密度は以下の式で表される．

$$f(\theta|x)\propto f(\theta)\cdot L(x|\theta) \qquad (35)$$

これを対数変換し，式（7），式（34）を代入すれば，

$$\ln f(\theta|x)\propto \ln f(\theta)+\ln L(x|\theta)$$

$$\propto -\frac{1}{2}(\theta-\theta^d)^t\Sigma_d^{-1}(\theta-\theta^d)$$

$$+\sum_{n=1}^{N}\sum_{i=1}^{m}x_{in}\cdot\ln(s_{in}) \qquad (36)$$

である．なお，式（36）の右辺第2項において s_{mn} は（$1-s_{1n}-\cdots-s_{(m-1)n}$）に等しい．以上によって，式（36）の最大値を求めれば新しいパラメータ θ を得られる．式（36）の第1項は最小2乗法と同等な線形式であるが，第2項には選択確率 s_{in} が含まれ，これが θ に関する非線形項になっている．ただし，第2項は集計モデルの最尤推定と同等な形式になっていることがわかる．したがって，本方法では一端，非集計モデルによって求めたパラメータを集計モデルのパラメータ推定と同時に更新することにより，集計モデルの再現性の確保と非集計モデルからの乖離を最小に止めることとの両者の折り合いで，パラメータ更新していると解釈することができる．

（3） 同時尤度最大化手法との関係

以上に提案した方法は同時尤度の最大化によっても定式化し，パラメータ推定を行うことができる．正規分布を仮定した従来の方法であれば，共役関数であることからベイズ更新法を用いる展開上の利点があった．

しかし，ディリクレ分布法では，式（30）の替わりに，

式（5）の選択確率 P_{in} を用いて，

$$L=\sum_{n=1}^{N_d}\sum_{i=1}^{m}\delta_{in}\cdot\ln P_{in}$$

$$+\sum_{i=1}^{m}(q_i-1)\ln z_i-\sum_{i=1}^{m}\Gamma(q_i) \quad\cdots\cdots\cdots\cdots\cdots\cdots\cdots (37)$$

なる同時尤度を考えることができる．ここで，δ_{in} は個人 n が選択した i のみで 1 を取るクロネッカーのデルタである．N_d は非集計データのサンプル数である．

一方，多項分布法では同様に式（5）の P_{in} を用いて，

$$L=\sum_{n=1}^{N_d}\sum_{i=1}^{m}\delta_{in}\cdot\ln P_{in}+\sum_{n=1}^{N}\sum_{i=1}^{m}x_{in}\cdot\ln s_{in} \quad\cdots\cdots\cdots\cdots\cdots (38)$$

を得る．この式では非集計モデルと集計モデルとの２つから共通のパラメータを求めることが直接行われている様子が明らかである．

以上の様に本研究で提案した方法では，同時尤度最大化によってもパラメータ更新を行えることを示した．この様な方法とベイズ更新型の推定方法との実用上の考察については後に述べることとする．

4. 各種更新方法の特性と予測への適用性

（1） 各種方法の特性の整理

本節では本研究で提案した各種方法の特性を整理する．既に筆者らが提案した正規分布を用いた方法も併せて，本研究で示した方法を列挙すれば，表—1 のようになる．

表—1　ベイズ更新法の種別の一覧

集計データの分布	方法の名称
正規分布	正規分布（逐次計算）法
	正規分布（線形近似）法
ガンマ分布	ディリクレ分布法
ポアソン分布	ﾎﾟｱｿﾝｰﾃﾞｨﾘｸﾚ分布法
多項分布	集計モデル統合法

ここで，正規分布（逐次計算）法は，非集計モデルの集計化のための関数式が非線形であることを考慮して，ベイズ更新を非線形関数に対して行うために逐次計算を要した．この方法では，集計データの誤差を分散として任意に導入できる大きな利点を有する．

一方，正規分布（線形近似）法は一般化最小 2乗法と等しく，逐次計算を必要とせずに更新解を求められる簡便な方法である．しかし，集計化の関数を線形近似するために元のパラメータまわりでテイラー展開を行う．したがって，パラメータの変化率が大きい場合には適用できない近似解法である．

ディリクレ分布法は，個々の集計データは独立に分布すると考えるが，そのシェアの分布を明示的に与えることによって，選択肢間の競合を考慮した更新を行える方法である．集計データの期待値に等しい分散を仮定するため，比較的小さな交通量が得られている場合に適すると考えられる．

また，ポアソンーディリクレ分布法は，交通量が離散量であることから，ガンマ分布をポアソン分布に置き直して整数のみを扱って更新を行う方法である．逐次計算の途上で毎回整数に置き直すため収束の安定性に欠けると考えられるが，ポアソン分布を交通量に仮定することから，ディリクレ分布法と同様に小さな交通量を用いて更新する場合に適すると考えられる．

表—2　非集計モデルと集計データ（交通量データ）とを統合する方法の一覧

更新法	推定法	交通量の分布	シェアの分布	交通量の誤差	適用範囲／備考
正規分布法	ベイズ推定 逐次推定	正規分布	―	分散共分散を設定	個々の交通量大
正規分布法（近似法）	ベイズ推定 GLS	正規分布	―	分散共分散を設定	個々の交通量大 ＊近似解法である
ディリクレ分布法	ベイズ推定 逐次推定	ガンマ分布	ディリクレ分布	期待値に等しい分散	比較的交通量小
ポアソンーディリクレ分布法	ベイズ推定 逐次推定	ポアソン分布	ディリクレ分布	期待値に等しい分散	個々の交通量小
多項分布法	ベイズ推定 逐次推定	多項分布	―	確定値	集計モデルの最尤推定との統合形
移転方法	最尤推定（α、β）	―	―	考慮せず	ｽｹｰﾙと定数項の再推定を行う方法
WESML法	重み付き最尤推定	―	―	考慮せず	集計シェアを重みを用いて選択肢別標本からﾊﾟﾗﾒｰﾀ推定

最後の多項分布法は，交通量の分布を多項分布に当てはめ，各選択肢の生起確率がシェアモデルによって構造化されていると考える方法である．最大化を図る式を見れば分かるように，非集計モデルより得たパラメータを集計モデルのパラメータの最尤推定によって修正する方法となっており，ベイズ更新法の中では最も直接的に，非集計モデルと集計モデルとの統合推定を行う形になっていることが分かる．

以上に示した各種方法の特徴を整理すると表—2 のようになる．追加的に得られるデータの性質によって，適宜使い分けを行える様子も明らかである．

（2） 予測への適用可能性

調査，推定，現況再現，移転，予測の各段階で，どの様な場面に本方法を活用できるか簡単に示す．非集計データと集計データとの統合利用が行えれば，調査段階においては家庭訪問調査や選択肢別調査とは別に，交通量調査や入り込み調査等を実施して，集計データを同時に集めておくことの意義を高めることが出来る．家庭訪問調査の数を減らしても交通量を計測することの有利さが示される場合も有り得る．

また，モデルのパラメータ推定段階においては本研究の方法を用いることにより，少数データの限界を直接的に補うことが可能になる．また，現況の交通流動を把握するためには，総量制約下でのゾーン細分化等が必要になるが，これにも非集計モデルと集計量との組み合わせ利用が有効である可能性がある．

さらに，モデルを他の地域へ移転する場合に新しい交通量データを用いて，元のモデルを修正移転する方法は，従来から最も オーソドックスな適用対象であり，より合

理的な方法があれば有意義であることも明らかである．また，将来予測において中間年のモデル更新のために数少ないODペアの交通量のみ再度調査して，これよりモデルパラメータの更新が行えるのであれば，効率的な予測の軌道修正が可能になる．以上のように，非集計モデルと集計データとの統合利用が予測作業に及ぼす効果は様々な場面にある．これらのうち定性的に自明である事柄は別にして，実証的な分析を進めなければ確認できないものについては今後の検討が期待される．

5. 新たなモデル更新法の適用事例

なお，以上の方法の基本的な特性については，簡単な実証計算によっても確認することができる．以下ではモデル更新に要する非線形最適化計算を実際の交通調査データに当てはめて行い，いくつかの特性を概観することとした．

（1） 適用データとモデル推定・更新結果の概要

平成2年11月に東京都・神奈川県で実施した「自動車による観光レジャー交通」に関する家庭訪問調査より得た557票を用いて非集計経路選択モデルを作成した．この調査では過去数カ月間に行った観光レジャーの旅行経験を詳細にアンケートしている．特に自由度の高い旅行における周遊・経路選択行動を正確にデータ化するために観光地域の地図を用いて，経路，宿泊地などを直接記入させる方式を採用した．これより，選択肢として高速道路と一般道路との2つを抜き出してデータ化し二肢選択のロジットモデルを作成した結果が表—3である．同表には222票の追加データを用いたベイズ更新の結果も併せて示してある．総所要時間はどの方法でも2倍程度に増しているが，総費用は逆に1桁小さく更新されている．経路ごとの海沿いの距離のパラメータはあまり大きく変化していないが，元々小さな定数項が大きく変化している様子も明らかである．計算にはGAUSS(Ver. 2.01）を利用し，収束アルゴリズムとしてBFGSおよびDFPを用いている．

なお，更新に用いた追加集計データには同年秋に実施された休日道路交通情勢調査を用いている．道路情勢調査において休日調査は昭和49年以来16年ぶりに再開され，全国の自動車所有者への大規模調査が平日調査に合わせて実施された．これよりゾーン間の分布交通量が別途拡大推計されているが，ここでは一部だけを取り出して本方法の適用イメージを明らかにするため，あえて表—4に示した発着地間の拡大前交通量を用いている．なお，拡大後交通量は正規分布法の追加データとして用いている．この場合，追加集計データの分散が必要になるが，ここでは以下の計算式からサンプリング誤差を求め分散の大きさとした．すなわち，ODごとに，高速道路利用率をP，サンプルOD交通量をn，拡大OD交通量

表—3 非集計モデルのパラメータ推定とベイズ更新の結果

	非集計ﾓﾃﾞﾙ	多項分布	ﾃﾞｨﾘｸﾚ分布	正規分布
総時間(分)	-0.00844 (2.13)	-0.0166	-0.0158	-0.0153
総費用(円)	-0.00201 (5.44)	-0.000154	-0.000133	-0.000128
海沿い距離(Km)	0.0170 (2.44)	0.0124	0.0144	0.0208
定数項(高速道路)	-0.0888 (0.462)	-0.896	-0.926	-0.972
尤度比 サンプル数	0.411 557	－	－	－

（　）内はt値

表—4 更新に用いた集計データとモデルによる推計結果

交通量の発生集中量	観測値		非集計ﾓﾃﾞﾙ		多項分布		ﾃﾞｨﾘｸﾚ分布		正規分布	
	高速	一般	高速	一般	高速	一般	高速	一般	高速	一般
東京A－箱根	3	2	0	5	3	2	3	2	3	2
川崎A	4	1	1	4	3	2	3	2	2	3
横浜A	1	12	2	11	5	8	5	8	4	9
横浜B	4	4	2	6	4	4	4	4	4	4
東京B－千葉	7	7	1	13	5	9	5	9	5	9
東京C	3	4	1	6	4	3	4	3	3	4
東京D	4	22	11	15	10	16	10	16	9	17
東京E－成田	4	2	0	6	3	3	3	3	3	3
東京F	4	1	1	4	2	3	2	3	2	3
東京C－勝浦	2	1	0	3	2	1	2	1	2	1
川口　－日光	3	2	0	5	4	1	4	1	3	2
東京G－東松山	5	3	1	7	3	5	3	5	3	5
八王子－相模湖	1	4	1	4	2	3	1	4	1	4
横浜C－御殿場	2	1	0	3	2	1	2	1	1	2
秦野	1	1	1	1	1	1	1	1	1	1
箱根　－東京H	2	1	0	3	1	2	1	2	1	2
東京A	3	1	0	4	1	3	1	3	1	3
川崎A	3	1	0	4	1	3	1	3	1	3
横浜A	1	11	1	11	5	7	4	8	4	8
横浜B	3	2	0	5	1	4	1	4	1	4
千葉　－東京B	6	4	1	9	4	6	4	6	4	6
東京C	3	2	1	4	3	2	3	2	2	3
東京D	7	25	14	18	13	19	12	20	12	20
成田　－東京E	1	2	0	3	2	1	1	2	1	2
東京F	3	1	0	4	2	2	2	2	2	2
伊香保－東京I	3	2	0	5	4	1	4	1	4	1
東松山－東京G	4	4	1	7	3	5	3	5	3	5
相模湖－八王子	1	4	1	4	2	3	1	4	1	4
つくば－東京D	4	3	1	6	4	3	4	3	3	4

高速：高速道路，一般：一般道路

をNとすれば，分散を，

$$V=\frac{N-n}{N-1}\cdot\frac{N^2\cdot P(1-P)}{n} \qquad (39)$$

によって表すことができる．他にも様々な誤差が発生していると考えられるが，ここでは上式のVによって追加データの信頼性が表されると考えている．

表—4にはモデル更新によって追加交通量の推計値が非集計モデルの推計結果と比べて変化する様子が示されている．当然ながら，左端の観測交通量に合わせるようにパラメータを更新しているため，非集計モデルによる再現結果よりも，観測値により近い結果が得られていることがわかる．多項分布法，ディリクレ分布法，正規分布法の3者で大きな差は見られないが，正規分布法と多項分布法とは若干他の関係と異なる傾向を見せている．なお，表中の発生集中地の詳細は付表—1に示した．また，より広範囲のデータを用いて目的地選択モデルなど

のパラメータ更新に本方法を用いることも可能である.

（2） 各更新法の基本特性の考察

次に追加する交通量の大きさや，元のモデルのパラメータ分散の大きさ等による更新結果の変動特性を簡単に調べた．図—2~5には表—4に既に示した追加交通量が仮にもっと大きな値であったと考えて，更新計算を行った結果を図示している．ここで言う大きな値とは，より高い抽出率で調査を行った場合を指し，具体的には交通量を a 倍（$a=1\sim30$）して，抽出率の増加を表現した．

交通量の増大は追加データ側をより良く推計するように働くため，結果としてパラメータが大きく変化すると想定される．ここでは，単純にサンプルを整数倍しているので，実際に抽出率を上げて調査した場合とは異なるが，追加データの増加によって各方法で同様に大きくパラメータが変わることを確認できた．なお，総費用のパラメータはどの方法でも抽出率が増すと正の値に変わり，効用関数の符号条件を満たさない修正となっている．また，変化の大きさが所要時間を除く3変数では正規分布，ディリクレ分布，多項分布の順，所要時間ではこの逆の順に大きくなっており，方法間で幾分差があることも明らかになった．

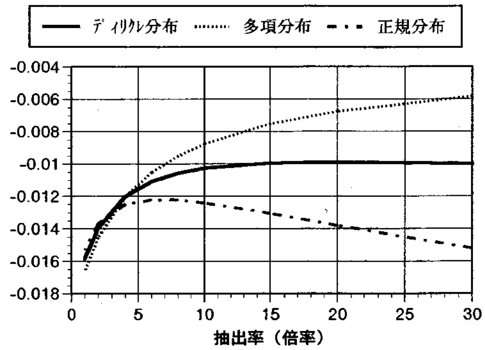
図—6には更新されたモデルを用いて，追加データに対する説明力を，交通量の誤差の絶対値和によって表してみた．これより，抽出率が元の10倍ほどになれば，集計量の残差で見る限り変化がほとんどなくなることがわかる．正規分布とディリクレ分布ではあまり差がなく，多項分布が少し大きな残差を持つこともわかる．

（3） 同時推定法との比較結果

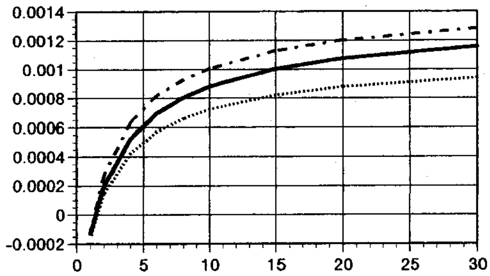
次に表—5にはディリクレ分布法と多項分布法とに対して，式（37）および式（38）の同時尤度最大化法を用いた場合とベイズ更新法との比較結果が示してある．本研究で同時推定を行わない理由の1つにパラメータ推定の安定性の問題を挙げているが，この点がどの程度異なるかを簡単に検討した．表には元の非集計データから222サンプルを10回繰り返して無作為に抽出し，それを追加データと同時に用いて各方法で推定したパラメータの分散とベイズ更新法によるパラメータの分散との比を示してある．

これより明らかなように，段階推定であるベイズ更新法では，本計算の様な2肢選択，4変数という簡単なケースであるにもかかわらず，推定結果の変動が同時推定の2割から4割ほど小さく済んでいることが明らかになった．このように変動が小さいことは，追加的に得られるデータの信頼性に問題がある場合にも，ベイズ更新法によればパラメータ修正を相対的に小さく安全側に済ませられることを示している．

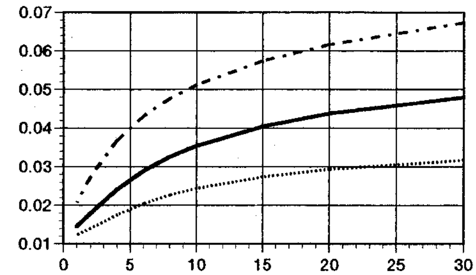
また以上の検討から，正規分布法に比べてディリクレ



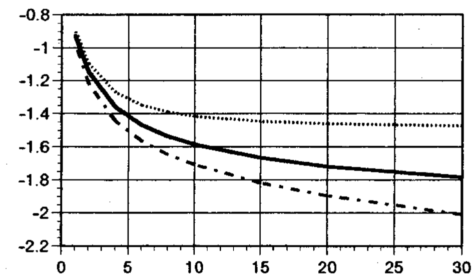
図—2 総所要時間パラメータの変化



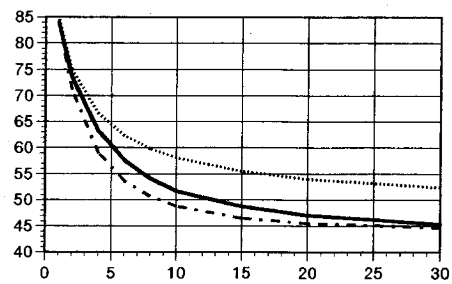
図—3 総費用パラメータの変化



図—4 海沿い距離パラメータの変化



図—5 定数項（高速道路）の変化



図—6 交通量の誤差の変化

表—5　パラメータ分散比
（ベイズ更新法／同時尤度最大化法）

説明変数＼推定法	ディリクレ分布法	集計モデル統合法
総所要時間	0．57	0．57
総費用	0．38	0．44
海沿い距離	0．82	0．83
定数項	0．78	0．79

付表—1

ｿﾞｰﾝ名		ｿﾞｰﾝ名	
東京A	新宿区、渋谷区	東京H	世田谷区、杉並区、大田区
B	江東区、墨田区	I	杉並区、練馬区
C	大田区、世田谷区	川崎A	中原区、高津区
D	足立区、葛飾区、江戸川区	横浜A	神奈川区、西区、保土ヶ谷区
E	文京区、中央区、千代田区	B	旭区、瀬谷区、緑区
F	江戸川区	C	神奈川区、港北区、緑区
G	北区、板橋区、練馬区		

分布法や多項分布法の修正結果が相対的に大きくないことが示されたが，この点では，従来実用的に用いてきた正規分布法に加え，少数データに対するこれら２つの方法の有効性が確認できたと考える．本計算例のようにデータが少ない場合にも，そのデータの持つ精度を反映した比較的穏やかな更新が行われていることを確認できたからである．

6.　おわりに

本研究では，データ統合によるモデル推定方法の体系的整理を行い，その後，従来から筆者らが提案していたベイズ更新方法を集計データに対する異なる仮定に基づいて再定式化し，それらの適用性を簡単に分析した．その結果，以下の２つを主な結論として挙げることができる．①種々の方法について適用条件を明らかにし，方法論の位置づけ，適用領域を明確にしたこと，②計算事例を通して定性的には見られない個々の方法の特性を概略把握できたことの２つである．なお，本研究で新たに提案した方法の実用性については，今後さらに検討を進めて明らかにする必要がある．ただし，更新の効果を当然期待できる大きな追加データの存在する場合のみならず，少数サンプルしか追加できないデータ環境で更新を行いたい場合にも，それに対応する適切な方法が存在することは大変に有意義である．将来の様々な交通調査，予測，政策分析を取り巻く環境変化に応じた分析方法論の幅広い展開が本研究に関わる今後の課題である．

最後に，本方法の展開にあたり有益なご示唆を頂いた東京工業大学工学部の繁桝算男先生，研究遂行にあたり協力を頂いた同大学同学部の岩倉成志助手，大学院生高田和幸氏に心より感謝いたします．

参　考　文　献

1）Atherton T. & M. Ben-Akiva : Transferability and Updating of Disaggregate Travel Demand Models, Transportation Research Record 610, pp.12～18, 1976.

2）Lerman S. & C. Manski : Sample Design for Discrete Choice Analysis of Travel Behavior, Transportation Research A, Vol.13, pp.29～44, 1979.

3）Ben-Akiva, M. & S. Lerman : Discrete Choice Analysis : Theory an Application to Travel Demand, MIT Press, 1985.

4）森地　茂・屋井鉄雄・平井節生：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，土木計画学研究論文集，No. 5，pp. 51～58，1987.

5）森地・屋井ほか：札幌市における地下鉄計画路線の需要予測調査，（財）運輸経済研究センター，1987.

6）Quevedo, S. : Combining Survey and Aggregate Data for Model Estimation, 1985.

7）Maher, M. : Inferences on Trip Matrices from Observations on Link Volumes ; a Bayesian Statistical Approach, Transportation Research B, Vol.17, No.6, pp.435～447, 1983.

8）Fisk, C.S. & D.E. Boyce : A Note on Trip Matrix Estimation from Link Traffic Count Data, Transportation Research B, Vol.17, No.3, pp.245～250, 1983.

9）Ben-Akiva, M. & T. Morikawa : Estimation of Switching Models from Revealed Preferences and Stated Intentions, Transportation Research A, Vol.24, No.6, pp.485～495, 1990.

10）Van Zuylen, H.J. & L.G. Willumsen : The Most Likely Trip Matrix Estimated from Traffic Counts, Transportation Research B, Vol.14, pp.281～294, 1980.

11）McNeil, S. & C. Hendrickson : A Note on Alternative Matrix Entry Estimation Techniques, Transportation Res., Vol.19B, pp.509～519, 1985.

12）Morichi, S. & T. Yai : Prediction of Trip Distribution by Disaggregate Behavioural Model, Proceedings of The World Conference on Transport Research, Vol.2, pp.1582～1599, 1986.

13）Ben-Akiva, M. : Methods to Combine Different Data Sources and Estimate Origin-Destination Matrices, Proc. of the 10th International Symposium on Transportation and Traffic Theory, pp.459～481, 1987.

14）Ben-Akiva, M. & T. Morikawa : Estimation of Travel Demand Models from Multiple Data Sources, In Transportation and Traffic Theory, M. Koshi, ed. pp.461～476, 1990.

15）屋井鉄雄・森地　茂：2，3の確率分布を用いた離散選択モデルの事後更新方法，東京工大土木工学科研究報告，No.44，pp.15～21，1991.

16）T. Yai : Disaggregate Behavioural Models and Their Applications in Japan, Transportation Research A, Vol.23, No.1 pp.45～51, 1989.

17）森川高行・山田菊子：SP データと RP データを用いた都市間鉄道のサービス改善に伴う需要予測法，土木計画学研究・講演集，No. 13，pp. 659～666，1990.

18）Morichi, S. & T. Yai : Estimation of Disaggregate Model using Additional Aggregate Data, Proceedings of PTRC Summer Annual Meeting, Bath, 1988.

19）屋井鉄雄，岩倉成志，魚谷　憲：多種データの複合調査による離散選択モデルの作成，日本行動計量学会第 20 回大会発表論文抄録集，pp.210～213，1992.

20）T.Yai, S. Moriti and K. Fan : Disaggregate Modeling by

Multi-Source Data, Proceedings of World Conference on Transport Research, 1990.

(1993. 2. 5 受付)

TRANSPORT DEMAND MODELS USING BAYESIAN UPDATING TECHNIQUE FOR COMBINED DATA SOURCES

Tetsuo YAI, Shigeru MORICHI and Satoshi UOTANI

This paper presented a new technique to estimate discrete choice model by individual and aggregated data sources. Bayesian updating approach was employed to improve the model predictability. Different methods using some kind of distributions : normal, multimonial and diricret were proposed here.
After the characteristics of these different methods were briefly investigated, an application for route choice behavior of car drivers had been demonstrated to provide the efficiency of the above methods. The results indicated that every method had appropriate data conditions.