

系列相関を持つRPデータとSPデータを同時に用いた離散型選択モデルの推定法

森川高行*・山田菊子**

信頼性の高いRPデータと操作性の高いSPデータを同時に用いて非集計行動モデルを推定することによって、両方のデータの補完的な長所を利用する方法論を著者らは提案してきた。本論文は、この方法論の発展形として、RPデータとSPデータの「系列相関」と「状態依存」を明示的に考慮したモデリングを提案し、交通機関選択分析に適用するものである。

Key Words : discrete choice model, RP and SP data, serial correlation, mode choice analysis

1. はじめに

非集計タイプの交通行動モデルを推定する際に用いられるデータには、大分して実際の行動に基づくRP(Revealed Preference) データと、実験者により提示される仮想の状況における被験者の選好の意思表示であるSP(Stated Preference) データがある。RPデータは実際の行動結果に基づいているため高い信頼性を有する反面、選択に影響を与える要因が不明瞭であったり、要因間の相関の強さのために効用関数のパラメータ推定に問題を生ずるなどの短所が指摘されている¹⁾。

一方、SPデータは一種の実験データであるため、現存しない代替案の選好を取り扱える上、要因の値を実験者が制御できるので、被験者が持つ要因間のトレードオフ関係(相対的重要度)を把握しやすいという長所を持つ。近年ではパーソナルコンピュータを用いて動的に条件を変える「応答型意識調査法」も適用されており^{2)~4)}、被験者に与える負担を軽減しつつ、多くの選好データを各個人から得ることも可能になっている。また、SPデータはその質問の形式によって選択データ、順位付けデータ、評点データなど様々な形態のものを得ることができる(RPデータとSPデータの性質のまとめは文献1)に詳しい)。

著者らは、互いに補完的な性質を持つRPデータとSPデータを同時に用いて、両データの長所を助長する非集計行動モデル推定の方法論を提案してきた^{5)~8)}。非集計行動モデルにおいてRPデータとSPデータを同時に用いる目的は、1)RPデータだけでは同定できない新規のサービスに対するパラメータを推定す

ること、2)とくにSPデータから得られる信頼性の高い、要因間のトレードオフに関する情報を用いてパラメータの信頼性(統計的有効性)をあげること、3)SPデータに含まれる様々なバイアスを修正すること、の3点に集約される。著者らがこれまでに提案した手法は、RPデータ、SPデータの発生過程をそれぞれランダム効用モデルとして定式化し、そのパラメータ推定をRPデータとSPデータを同時に用いて行なうことであった。ここではそれぞれRPモデルとSPモデルで表される両データの発生過程に含まれる誤差が、同一個人内においても独立であるという仮定に基づいてパラメータ推定を行っていた。

この方法論を用いたいくつかの実証的研究によると、SP質問の回答が実際の行動に大きく影響を受けることが多いことがわかった^{9)~11)}。これは、過去の行動結果が現在の状態を形成し、その状態が将来の行動に影響を与えるという「状態依存」に近い性質と考えられる。これに対してこれまでの定式化では、SPモデルの効用関数の中に実際の行動結果を表すダミー変数を入れることで対処してきた。また、一部のケーススタディでは、SPデータとRPデータの独立性を仮定した定式化で推定されるパラメータの一部に明かなバイアスが現れた⁹⁾。これは同一個人内で、RPモデルおよび複数回答のSPモデルの効用の誤差項が独立でない「系列相関」と、先に述べた「状態依存性」が同時に存在するために生じたバイアスであると考えられる。

そこで本論文ではまず、RPモデルとSPモデルに存在する「系列相関」についてその内容を検討した上で、「系列相関」と「状態依存」という2つの性質を明示的に考慮するモデルの定式化および推定方法の提案を行う。提案した手法の有効性の検討は、比較的簡便な設計の調査票によって得られた日本での

* 正会員 Ph.D. 名古屋大学助教授 工学部土木工学科 (〒464-01 名古屋市中千種区不老町)

** 正会員 工修 (株)三菱総合研究所 社会システム第二部

調査データと、応答型意識調査によるオランダでの調査データの2種類について行った。両調査の対象はともに2つの地域を結ぶ都市間交通における交通機関選択行動である。

2. RPデータとSPデータの系列相関

確率的選択モデルの中でも、ランダム効用モデルの基本的考え方は、以下のとおりである。意思決定者は自分にとって最大の効用をもたらす選択肢を選択するが、分析者には、意思決定者の効用に影響を与える要因値と、それが効用という一次元数にどの程度影響を与えるかを完全には観測できないため、効用を(分析者にとっては)確率変数にせざるを得ない。効用を確定部分と確率部分に分けられるとすると、効用の確定部分を与えたときの意思決定者の行動は、観測者にとって確率的選択となる。このとき、効用の確定項を通常のように、属性値と重み(係数)の線形和で表すと、確率項(誤差項)には、選択肢や意思決定者に関する観測できなかった属性、係数で表される「嗜好」のばらつき、属性値の測定誤差や関数形の誤りなどが含まれる。

例えば、通常の交通機関選択モデルでは、旅行時間や費用、乗り換え回数、意思決定者の性別といった計測が容易な属性のみを効用関数の確定項の説明変数に用いている。このようなモデルでは、計測が困難であるために説明変数に加えられなかった要因が誤差項に多く含まれると考えられる。つまり、快適性や安全性など主観的な属性や、「車の運転が好きである」とか「費用は全く問わない」といった潜在的な態度は、誤差項にすべて入ってしまう。このような要因は、RPデータ、SPデータに関わらずその個人の嗜好に大きく影響を与えることが多く、通常の説明変数で表現することは困難なものの、モデル構築の際になんらかの形で考慮されることが望ましい。本論文では、このような要因が選択肢および意思決定者ごとに特有なものであることから、「個人・選択肢に共通なシステムティックな誤差」と呼ぶことにする。

システムティックな誤差が存在する場合、同一個人から取ったRPデータとSPデータの生成過程を、それぞれランダム効用モデルでRPモデル、SPモデルと表したときには、それぞれのモデルの誤差項間で、同一個人内の、とくに同一選択肢間で強い相関が生じる。RPモデル、SPモデルの効用関数の誤差項間の相関は、一種の系列相関(serial correlation)と考えられる。系列相関を考慮せず、RPデータとSPデータをプールして取り扱うこれまでに提案した手法の場合、系列相関を考慮しないことによりパラメータの推定

値は漸近的一致性は保たれるが有効性が失われることが知られている⁹⁾。

RPデータとSPデータを同時に用いる本手法でとくに重大な問題となるのは、系列相関と同時に状態依存(state dependence)が存在するために生じるパラメータ推定値のバイアスである。動的分析では系列相関を「見せかけの状態依存」、状態依存を「真の状態依存」と呼び明確な区別を与え、これらを的確に取り扱う方法が示されている¹⁰⁻¹²⁾。これまでに提案したRPデータとSPデータを同時に用いる選択行動モデルでは「真の状態依存」としてSPモデルの説明変数にRPモデルの選択結果を取り込むことで対処しているが、「見せかけの状態依存」については、RPモデルとSPモデルの誤差項を互いに独立であると仮定しているため全く考慮していない。この両者が存在したときは、誤差項と説明変数が相関を持ち、パラメータの推定値に重大なバイアスが生じている可能性があった。

3. RP/SP同時推定法の概要

本章ではこれまでに著者らが提案したRPデータとSPデータを用いた選択行動モデルの同時推定法(RP/SP同時推定法)の概略について述べる。(より詳しい説明については、文献5)~8)または13)を参照されたい。)

以下に示すRP/SP同時推定法のフレームワークでは、RPデータを発生するRPモデルとSPデータを発生するSPモデルの2つを考え、同一の個人に対しては主要な属性間のトレードオフ関係はRPモデルとSPモデルで共通と考えている。

RPモデル

$$U_{in}^{RP} = \beta' x_{in}^{RP} + \alpha' w_{in}^{RP} + \varepsilon_{in}^{RP} \\ = V_{in}^{RP} + \varepsilon_{in}^{RP}, \quad i = 1, \dots, I_n^{RP} \text{ and } n = 1, \dots, N^{RP} \quad (1)$$

$$d_n^{RP}(i) = \begin{cases} 1: & \text{if individual } n \text{ chose alternative } i \text{ in the RP data} \\ 0: & \text{otherwise} \end{cases}$$

SPモデル

$$U_{in}^{SP} = \beta' x_{in}^{SP} + \gamma' z_{in}^{SP} + \varepsilon_{in}^{SP} \\ = V_{in}^{SP} + \varepsilon_{in}^{SP}, \quad i = 1, \dots, I_n^{SP} \text{ and } n = 1, \dots, N^{SP} \quad (2)$$

$$d_n^{SP}(i) = \begin{cases} 1: & \text{if individual } n \text{ chose alternative } i \text{ in the SP data} \\ 0: & \text{otherwise} \end{cases}$$

ただし、

U_{in} = 個人 n の選択肢 i に対する総効用

V_{in} = 総効用 U_{in} の確定項

ε_{in} = 総効用 U_{in} の誤差項

$d_n(i)$ = 個人 n の選択肢 i に対する選択結果

x_{in}, w_{in}, z_{in} = 個人 n の選択肢 i の確定効用の説明変数ベクトル

α, β, γ = 未知パラメータベクトル

上記の定式化のSPデータは、提示されたいくつかの選択肢の中から最も望ましいものを答える、いわゆる「選択データ」として理論展開を行なっているが、その他の回答形式、例えば順位付けデータであっても、選択確率と未知パラメータ推定の際の尤度関数の表現が異なるだけで理論の一般性を失うものではない。

この定式化において、 x はRPモデルとSPモデルで共通の係数ベクトル β を持つ説明変数ベクトルであり、 w, z はそれぞれRPモデル、SPモデルで異なる係数を持つ説明変数ベクトルである。つまり、 $\gamma'z$ がSPバイアスおよびSPデータにしか含まれ得ない属性項（例えば新しいサービスの影響）を表しており、これによってSPデータに特有のバイアスを除去することができる。例えば、「過去の行動の正当化」(実際の選択結果を正当化するようにSP質問においても選択すること)や「政策操縦」(SPデータが政策決定に影響を与える場合、自分にとって好ましい政策を誘導するようにSP質問においても選択すること)のようなSPバイアスは、SPモデル固有の定数項に含まれると考えられるため、RPモデルとSPモデルで別の定数項を推定し(つまり、選択肢を表すダミー変数をそれぞれのモデルに固有なベクトル w, z の一部として定式化する)、需要予測に用いる効用関数においては、RPモデルの定数項を利用することが考えられる(詳しくは文献13)参照)。

RPデータ、SPデータそれぞれを用いて式(1)、(2)の未知パラメータを別々に推定すると、効用関数の誤差項のばらつき大きさに従ったスケールを持つ推定値が得られる。そこで両モデルのスケールを統一するためにスケールパラメータを導入する。両モデルの効用関数のスケールを同一にすることにより、推定の際にそれぞれの係数を両モデルで共有することができる。スケールパラメータ μ は、式(3)のようにそれぞれのモデルの誤差項の分散の比で表現される。このスケールパラメータの推定値が1よりも小さければ、SPモデルではRPモデルよりも誤差項が卓越していることになる。

$$\text{Var}(e_{in}^{RP}) = \mu^2 \text{Var}(e_{in}^{SP}), \quad \forall i, n \quad \dots\dots\dots(3)$$

すべての未知パラメータは最尤推定法を用いて推定される。各モデルの誤差項が独立であるという仮定に基づき尤度関数を定義すれば、RPモデルの対数尤度関数(式(4))、SPモデルの対数尤度関数(式(5))の和(式(6))がRP/SP同時推定モデルの対数尤度関数となる。

$$L^{RP}(\alpha, \beta) = \sum_{n=1}^{N^{RP}} \sum_{i=1}^{J_n^{RP}} d_n^{RP}(i) \ln(P_n^{RP}(i)) \quad \dots\dots\dots(4)$$

$$L^{SP}(\beta, \gamma, \mu) = \sum_{n=1}^{N^{SP}} \sum_{i=1}^{J_n^{SP}} d_n^{SP}(i) \ln(P_n^{SP}(i)) \quad \dots\dots\dots(5)$$

$$L^{RP+SP}(\alpha, \beta, \gamma, \mu) = L^{RP}(\alpha, \beta) + L^{SP}(\beta, \gamma, \mu) \quad \dots\dots\dots(6)$$

以上の手法により求められるパラメータの推定量は、RPモデルとSPモデルの誤差項が独立な場合には一貫性および漸近的有効性を持ち、独立でない場合は一貫性を持つが、漸近的有効性を持たない推定量が得られる⁹⁾。

4. 系列相関を考慮したRP/SP同時推定法

本章では、2.で論じたRPデータとSPデータの系列相関を明示的に考慮したRP/SP同時推定モデルを提案する。通常選択モデルの効用関数の誤差項には、前述のように、個人・選択肢に特有な要因が卓越していることが多い。そこで、提案するモデルでは、誤差項を個人・選択肢に共通なシステムティックな部分と、分析者にとって真にランダムなホワイトノイズの部分に分解する。このとき、RP/SP同時推定モデルのフレームワークでは、システムティックな誤差項を両モデルで共有することになる。つまり、システムティックな誤差項を λ 、ホワイトノイズを表す部分を v とすると以下のような誤差項の分解が可能である。

$$e_{in}^{RP} = \lambda_{in} + v_{in}^{RP} \quad \dots\dots\dots(7)$$

$$e_{in}^{SP} = \theta_i \lambda_{in} + v_{in}^{SP} \quad \dots\dots\dots(8)$$

λ に付加された未知係数 θ は、以下のような役割を果たす。システムティックな誤差の効用に及ぼす影響は、RPモデルとSPモデルで同じである理由はない。また、 λ_i を選択肢について同一の分布とした場合、選択肢ごとのシステムティックな誤差の大きさの違いを $\theta_i \lambda_i$ によって表すことができる。

これより系列相関を考慮した場合のRP、SPそれぞれのモデルの効用関数は次のようになる。

$$U_{in}^{RP} = \beta' x_{in}^{RP} + \alpha' w_{in}^{RP} + \lambda_{in} + v_{in}^{RP} \\ = V_{in}^{RP} + \lambda_{in} + v_{in}^{RP}, \quad i = 1, \dots, J_n^{RP}, n = 1, \dots, N_n^{RP} \quad (9)$$

$$U_{in}^{SP} = \beta' x_{in}^{SP} + \gamma' z_{in}^{SP} + \theta_i \lambda_{in} + v_{in}^{SP} \\ = V_{in}^{SP} + \theta_i \lambda_{in} + v_{in}^{SP}, \quad i = 1, \dots, J_n^{SP}, n = 1, \dots, N_n^{SP} \quad \dots\dots\dots(10)$$

さらに前章で述べたようにRPデータとSPデータでホワイトノイズのスケールが同じである理由はないので、同様にスケールパラメータ μ を導入して、効

用関数のスケールを統一する。

$$\text{Var}(v_{in}^{\text{RP}}) = \mu^2 \text{Var}(v_{in}^{\text{SP}}) \dots\dots\dots (11)$$

RPモデル, SPモデルの選択確率は, これらの効用変数を用いて, まず各個人の選択肢ごとに共通な誤差項が既知であるときの条件付き確率として次のように表現される。

$$P_n^{\text{RP}}(i|\lambda) = \text{prob}(U_{in}^{\text{RP}} > U_{jn}^{\text{RP}}; i, j=1, \dots, J_n^{\text{RP}}, \text{ and } i \neq j | \lambda) \dots\dots\dots (12)$$

$$P_n^{\text{SP}}(i|\lambda) = \text{prob}(U_{in}^{\text{SP}} > U_{jn}^{\text{SP}}; i, j=1, \dots, J_n^{\text{SP}}, \text{ and } i \neq j | \lambda) \dots\dots\dots (13)$$

ただし, λ はシステマティックな誤差 λ_i を要素に持つ確率変数ベクトルである。

例えば v に, 独立で各データごとに同一のスケールを持つガンベル分布を仮定すれば, 式(12), (13)は通常の多項ロジットの選択確率で表される。

ある個人について, 実際の行動結果とひとつのSP質問に対する選択結果が得られている場合, それぞれについて, i, j という選択肢を選択する同時選択確率は, システマティックな誤差項 λ の確率密度関数 $f(\lambda)$ を与えることにより次のように表される。

$$P_n(d_n^{\text{RP}}(i) = 1, d_n^{\text{SP}}(j) = 1) = \iiint \dots \int_{-\infty}^{+\infty} P_n^{\text{RP}}(i|\lambda) P_n^{\text{SP}}(j|\lambda) f(\lambda) d\lambda_{1n} d\lambda_{2n} \dots d\lambda_{I_n n} \dots\dots\dots (14)$$

λ の分布形としては互いに独立な多変量正規分布などが考えられる。

効用関数中の係数 $\alpha, \beta, \gamma, \theta$, およびスケールパラメータ μ の推定は, サンプル内のすべての個人のRP/SP同時選択確率の同時生起確率である尤度関数の(対数の)最大化問題を解くことにより行う。

$$\max_{\alpha, \beta, \gamma, \theta, \mu} L(\alpha, \beta, \gamma, \theta, \mu) = \sum_{n=1}^N \sum_{i=1}^{J_n^{\text{RP}}} \sum_{j=1}^{J_n^{\text{SP}}} [(d_n^{\text{RP}}(i) \cdot d_n^{\text{SP}}(j)) \cdot \ln \{P_n(d_n^{\text{RP}}(i) = 1, d_n^{\text{SP}}(j) = 1)\}] \dots\dots\dots (15)$$

個人ごとの尤度の計算には, 式(14)で示されたように多重(選択肢の数だけ)の数値積分を必要とする。著者らの推定計算には, IBM-DOSパーソナルコンピュータ用の統計計算用言語GAUSS386を用いてプログラムしたが, 数値計算サブルーチンを用いても実用上計算時間は全く問題ないことがわかっていく(後述のケーススタディでもパソコンで30分以内)。以上が, 系列相関を考慮したRP/SP同時推定モデルである。SPデータが, 選択データでなく順位付けや一対比較の場合も, ランダム効用モデルを用いることによって同様に定式化できる。個人から複数のSP

データが得られているときは, 式(14)の同時選択確率に複数のSPモデルが入ることになる。

また, 状態依存性(SPデータの選択行動がRPデータの選択結果に依存する性質)については, RPデータの選択結果をSPモデルの説明変数(以下では慣性項と称する)として用いることで対応できる。

$$U_{in}^{\text{SP}} = V_{in}^{\text{SP}} + \gamma \delta_{in}^{\text{RP}} + \varepsilon_{in}^{\text{SP}} \dots\dots\dots (16)$$

$$\delta_{in}^{\text{RP}} = \begin{cases} 1: \text{if individual } n \text{ chose alternative } i \text{ in the RP data} \\ 0: \text{otherwise} \end{cases}$$

以上述べた推定法によって得られるパラメータ推定量は, 誤差項の分解が仮定に従うかぎり, 漸近的に有効で一致性を持つ。

5. ケーススタディ

(1) データの概要

前章までで提案したモデルの実証的研究として, 表一に示す2つのデータを用いてケーススタディを行なった。両データともに都市間交通における機関選択データで, 各個人からRPデータとSPデータを取っているが, 2つのデータの大きな違いはSPデータの形式にある。データAは各交通機関ごとに全部で3種類のSP質問が提示されているのに対し, データBはラップトップ・コンピュータを用いた応答型意識調査が適用され, 回答者ごとにRPデータの属性値を考慮して動的にシナリオが設定されているため, 実験計画的視点からはより精度の高いデータが得られていると考えられる。本ケーススタディでは, 提案した手法がこのように異なった性質を持つ2種類のデータに対してどれほど有効性を持つか検討する。なお, この2種類のデータに関する説明は文献(4), (14)に詳しい。

(2) データAによる推定結果

データAは, 鉄道, 高速バス, 自家用車の3項選択であるため, RPモデル, SPモデルのいずれについてもホワイトノイズを表す誤差項 v に, 独立でかつRPデータ, SPデータ内で同一な分散を持つガンベル分布を(正規分布の近似として)仮定するロジットタイプのモデルを用いた。

以下, 未知パラメータの推定結果を表二, 表三に示しその内容を検討する。

a) RPモデルとSPモデルの推定

まず, RPデータ, SPデータそれぞれを個別に用いたモデルを推定した結果が表二の左側2列である。RPモデル, SPモデルともに符号的には正しいパラメータの推定値が得られた。しかし, SPモデルについては, 慣性項が非常に強い影響力を持ち, 所要時間などの説明力が低くなった。

表-1 ケース・スタディに用いたデータの概要

| | データA | データB |
|--------|--|---|
| 調査目的 | ・2都市間を結ぶ高速交通機関 ・快適性の高い新型車両を用いた特急の増便による転換交通の予測 | ・ナイメゲン-ランドシユタット大都市圏間の高速交通機関の需要予測 |
| 代替交通機関 | ・鉄道(特急) ・高速バス ・自家用車 | ・鉄道 ・自家用車 |
| 調査法 | ・選択肢別抽出による調査票配布、郵送回収 | ・ランダムサンプリングによる住民調査 ・ラップトップコンピューターを用いた「応答型意識調査」 |
| サンプル数 | ・RP: 255 ・SP: 679 | ・RP: 228 ・SP: 1511 |
| 調査時期 | ・1990年2月(日本) | ・1987年(オランダ) |
| SPデータ | ・選択データ | ・対比較5段階評定データ |

b) RP/SP同時推定

3.で示した系列相関を考慮しないRP/SP同時推定モデルが表-2の右側2列である。効用関数の説明変数としては、SPモデルに固有なものに、RPデータの選択結果である慣性ダミーと快適性の高い列車を用いる新型特急の設定の有無を表す新型特急ダミー、両モデルに共通なものとして業務目的トリップの旅行費用、全目的の所要時間を用いた。また、定数項はそれぞれのモデル固有の値として推定した。

慣性ダミーを持たないモデルの推定値では、新型特急ダミーにかかる係数が負の推定値を示した。このダミー変数はSP質問において、料金は従来型特急と全く変わらないが所要時間が若干短く快適性の高い新型車両が用いられた場合のみ鉄道の効用関数に含まれる変数である。これは明らかに鉄道の効用を高めるものであり、その係数の符号は正の値を取るはずである。

負の値となった原因としては、SP質問において快適性の高い新型車両を用いた特急の増便を含むシナリオが提示されたのが高速バス、自家用車の利用者に対してのみであったことが考えられる。すなわちSPモデルの鉄道の効用関数で新型特急ダミーが1となるのは、現在鉄道を利用していない回答者に限られたため、あたかも「鉄道以外の交通機関利用ダミー」、あるいは「鉄道以外の交通機関利用の慣性ダミー」として、鉄道の効用を低下させる方向に働いたと思われる。これはとりもなおさず、自家用車や高速バスの効用関数の誤差項には、説明変数に用いられた要因以外にも鉄道を利用しないシステムティックな誤差要因が含まれ、このシステムティックな

表-2 データAによるパラメータ推定結果(1)
()内t値

| | RPモデル | | SPモデル | | RP/SP同時推定モデル | |
|-----------------|------------------|---------|-------------------|---------|------------------|------------------|
| | | | | | 慣性項なし | 慣性項あり |
| 鉄道定数項(RP) | 1.53 (4.2) | | | | 1.72 (5.3) | 1.67 (5.1) |
| バス定数項(RP) | 1.24 (4.7) | | | | 1.27 (4.8) | 1.26 (4.8) |
| 鉄道定数項(SP) | | | 1.83 (2.0) | | 1.00 (3.7) | 3.37 (3.3) |
| バス定数項(SP) | | | 0.586 (2.0) | | 0.147 (0.9) | 1.17 (1.7) |
| 業務目的旅行時間(時間) | -0.553 (-1.5) | | -0.0227 (-0.1) | | -0.256 (-1.5) | -0.339 (-1.4) |
| 1人当たり費用(1000円) | -0.354 (-3.8) | | -0.216 (-4.4) | | -0.389 (-4.4) | -0.383 (-4.4) |
| 新型特急ダミー | | | 0.768 (3.6) | | -0.296 (-1.3) | 1.49 (2.3) |
| バス慣性項 | | | 1.53 (5.4) | | | 2.87 (2.5) |
| 自家用車慣性項 | | | 3.19 (10.7) | | | 6.18 (2.8) |
| スケールパラメータ | | | | | 0.859 (3.9) | 0.518 (3.0) |
| N | 255 | 679 | 934 | 934 | | |
| L(0) | -191.35 | -515.65 | -707.00 | -707.00 | | |
| L(β) | -150.15 | -386.44 | -614.91 | -537.45 | | |
| ρ ² | 0.215 | 0.251 | 0.130 | 0.240 | | |
| ρ̄ ² | 0.194 | 0.237 | 0.119 | 0.226 | | |

部分λと、状態依存を表す「新型特急ダミー」が相関を持ったために、このダミー変数の係数に重大なバイアスが現われたと考えられる。

c) 系列相関を考慮したRP/SP同時推定

式(9)~(11)に従い、効用関数を以下のように定めた。

$$U_{Rail,n}^{RP} = \beta' x_{Rail,n}^{RP} + \alpha' w_{Rail,n}^{RP} + \lambda_{Rail,n} + v_{Rail,n}^{RP} \dots (17)$$

$$U_{Bus,n}^{RP} = \beta' x_{Bus,n}^{RP} + \alpha' w_{Bus,n}^{RP} + \lambda_{Bus,n} + v_{Bus,n}^{RP} \dots (18)$$

$$U_{Car,n}^{RP} = \beta' x_{Car,n}^{RP} + \alpha' w_{Car,n}^{RP} + v_{Car,n}^{RP} \dots (19)$$

$$U_{Rail,n}^{SP} = \beta' x_{Rail,n}^{SP} + \gamma' z_{Rail,n}^{SP} + \theta_{Rail} \lambda_{Rail,n} + v_{Rail,n}^{SP} \dots (20)$$

$$U_{Bus,n}^{SP} = \beta' x_{Bus,n}^{SP} + \gamma' z_{Bus,n}^{SP} + \theta_{Bus} \lambda_{Bus,n} + v_{Bus,n}^{SP} \dots (21)$$

$$U_{Car,n}^{SP} = \beta' x_{Car,n}^{SP} + \gamma' z_{Car,n}^{SP} + v_{Car,n}^{SP} \dots (22)$$

誤差項λが3つの代替案のうちの2つの効用関数にのみ含まれるのは、λが次に示す条件付き選択確率においてあたかも定数項のように働き、選択肢と同じ数だけ含めた場合、パラメータが同定できない

ためである。

ここで誤差項の分布形として、 v については前述のようなガンベル分布、 λ については各選択肢間で独立な標準正規分布を仮定した。まず、システムティックな誤差項 λ が既知であるという条件のもとでの各モデルの選択確率は、

$$P_n^{RP}(i|\lambda) = \frac{\exp(\beta'x_{in}^{RP} + \alpha'w_{in}^{RP} + \lambda_{in})}{\sum_{k=1}^{I_n^{RP}} \exp(\beta'x_{kn}^{RP} + \alpha'w_{kn}^{RP} + \lambda_{kn})} \dots\dots (23)$$

$$P_n^{SP}(j|\lambda) = \frac{\exp\{\mu \cdot (\beta'x_{jn}^{SP} + \gamma'z_{jn}^{SP} + \theta_j\lambda_{jn})\}}{\sum_{m=1}^{J_n^{SP}} \exp\{\mu \cdot (\beta'x_{mn}^{SP} + \gamma'z_{mn}^{SP} + \theta_m\lambda_{mn})\}} \dots\dots (24)$$

である。ここで、 λ の同時分布関数を用いることにより、RPモデルとSPモデルの同時選択確率を得る。

$$P_n\{d_n^{RP}(i) = 1, d_n^{SP}(j) = 1\} = \int \dots \int P_n^{RP}(i|\lambda_n) P_n^{SP}(j|\lambda_n) \phi(\lambda_{Bus,n}) \phi(\lambda_{Car,n}) d\lambda_{Bus,n} d\lambda_{Car,n} \dots\dots (25)$$

ここに $\phi(\lambda)$ は標準正規分布の密度関数である。

式(25)の計算には数値積分を要するが、前述のように計算時間上の問題はない。

表-3に未知パラメータの推定結果を示す。系列相関のみを考慮したモデルでは慣性項を含まなくとも新型特急ダミーは有意に正の値となり、前述のバイアスが改善された。また、スケールパラメータも系列相関を考慮しないモデルの場合に比べ1.0に近づいており、これはRPモデルとSPモデルのホワイトノイズのスケールが同じ程度になっていることを示している。系列相関に加え、慣性項を加えることで、スケールパラメータは、1.11とさらに1.0に近づき、新型特急ダミーもさらに有意な正値となった。また、系列相関を考慮しない場合に比べ、慣性項の影響が小さくなっている。系列相関を考慮したモデルはいずれも、考慮しないモデルに比べ、モデルの適合度は有意に高くなっている。

(3) データBによる推定結果

データBは2項選択および一対比較であり、RPモデル、SPモデルそれぞれに2項プロビットモデルを適用した。SPデータは一対比較で、1)必ず選択肢1を選択する、2)たぶん選択肢1を選択する、3)どちらともいえない、4)たぶん選択肢2を選択する、5)必ず選択肢2を選択する、といった5段階評定データであるため、本来であればオーダー・プロビット・モデルを用いるべきであるが、同時推定における計算を簡略にするため、「どちらともいえない」

表-3 データAによるパラメータ推定結果(2)
() 内 t 値

| | RP/SP同時推定モデル | | 系列相関を考慮したRP/SP同時推定モデル | |
|----------------|------------------|------------------|-----------------------|-------|
| | 慣性項あり | 慣性項なし | 慣性項なし | 慣性項あり |
| 鉄道定数項(RP) | 1.67 (5.1) | 1.90 (3.8) | 1.73 (5.1) | |
| バス定数項(RP) | 1.26 (4.8) | 1.59 (4.3) | 1.38 (4.0) | |
| 鉄道定数項(SP) | 3.37 (3.3) | 1.83 (1.53) | 2.01 (3.4) | |
| バス定数項(SP) | 1.17 (1.7) | 1.10 (1.1) | 1.33 (2.1) | |
| 業務目的交通時間(時間) | -0.339 (-1.4) | -0.225 (-0.4) | -0.508 (-1.5) | |
| 1人当たり費用(1000円) | -0.383 (-4.4) | -0.371 (-3.4) | -0.417 (-4.8) | |
| 新型特急ダミー | 1.49 (2.3) | 0.796 (2.6) | 1.23 (2.7) | |
| バス慣性項 | 2.87 (2.5) | | 0.0823 (0.2) | |
| 自家用車慣性項 | 6.18 (2.8) | | 3.46 (3.0) | |
| θ -鉄道 | | 2.41 (2.1) | 2.08 (3.7) | |
| θ -バス | | 2.56 (1.4) | 1.48 (2.8) | |
| スケールパラメータ | 0.518 (3.0) | 1.25 (3.1) | 1.11 (3.3) | |
| N | 934 | 255 | 255 | |
| $L(0)$ | -707.00 | -707.00 | -707.00 | |
| $L(\beta)$ | -537.45 | -527.66 | -507.30 | |
| ρ^2 | 0.240 | 0.254 | 0.282 | |
| $\bar{\rho}^2$ | 0.226 | 0.240 | 0.265 | |

と答えたサンプルを除いて選択データとして取り扱った。このように簡略化した2項プロビットとオーダー・プロビットの推定結果を比較しても、効用関数の係数の推定値は同様な値を持つことが確かめられた⁹⁾。

a) RPモデルとSPモデルの推定

RPデータ、SPデータそれぞれ独立に2項プロビットモデルを適用した結果が表-4の左側2列である。

交通機関の属性を示す4つの未知パラメータの推定値は両モデルともに負値となったが、SPモデルでは乗り換え回数については有意な値となっていない。また定数項および個人の属性を表す2つのパラメータについては両モデルで異なった符号が推定された。

表-4 データBによるパラメータ推定結果(1)
() 内 t 値

| | RPモデル | SPモデル | RP/SP同時推定モデル | |
|----------------------|-------------------|-------------------|-------------------|-------------------|
| | | | 慣性項なし | 慣性項あり |
| 鉄道定数項(RP) | 0.501 (2.1) | | 0.564 (2.9) | 0.658 (3.0) |
| 鉄道定数項(SP) | | -0.927 (-7.9) | -0.124 (-0.9) | -3.53 (-2.9) |
| 1人あたり費用 (Guilder) | -0.0270 (-3.9) | -0.0104 (-4.5) | -0.0118 (-2.9) | -0.0311 (-5.4) |
| 車内旅行時間 (時間) | -0.342 (-1.3) | -0.273 (-3.0) | 0.415 (4.0) | -0.523 (-2.6) |
| 車外旅行時間 (時間) | -1.61 (-3.8) | -0.203 (-1.3) | -0.185 (-5.1) | -1.48 (-3.9) |
| 乗り換え回数 | -0.139 (-0.9) | -0.0102 (-0.2) | -0.317 (-3.6) | -0.116 (-0.9) |
| 業務目的ダミー | 0.902 (3.4) | -0.111 (-0.1) | 0.896 (4.1) | 0.523 (1.9) |
| 女性ダミー | 0.188 (2.5) | -0.0635 (-0.8) | 0.301 (2.5) | 0.284 (1.5) |
| 慣性項 | | 1.65 (16.3) | | 5.59 (3.6) |
| スケールパラメータ | | | 0.670 (4.2) | 0.268 (3.6) |
| N | 228 | 1511 | 1739 | 1739 |
| L(0) | -158.04 | -1047.35 | -1205.38 | -1205.38 |
| L(β) | -112.56 | -655.94 | -937.90 | -774.25 |
| ρ ² | 0.288 | 0.373 | 0.222 | 0.358 |
| ρ̄ ² | 0.243 | 0.366 | 0.214 | 0.349 |

表-5 データBによるパラメータ推定結果(2)
() 内 t 値

| | RP/SP同時推定 モデル 慣性項あり | 系列相関を考慮したRP/SP 同時推定モデル | |
|----------------------|---------------------------|---------------------------|-------------------|
| | | 慣性項なし | 慣性項あり |
| 鉄道定数項(RP) | 0.658 (3.0) | 0.123 (0.5) | 0.297 (1.3) |
| 鉄道定数項(SP) | -3.53 (-2.9) | -0.323 (-0.9) | -0.823 (-2.1) |
| 1人あたり費用 (Guilder) | -0.0311 (-5.4) | -0.0331 (-5.6) | -0.0380 (-6.2) |
| 車内旅行時間 (時間) | -0.523 (-2.6) | -0.128 (-0.6) | -0.457 (-2.2) |
| 車外旅行時間 (時間) | -1.48 (-3.9) | -1.60 (-3.2) | -1.38 (-2.6) |
| 乗り換え回数 | -0.116 (-0.9) | -0.201 (-1.5) | -0.170 (-1.1) |
| 業務目的ダミー | 0.523 (1.9) | 0.877 (3.4) | 0.567 (2.1) |
| 女性ダミー | 0.284 (1.5) | 0.547 (2.6) | 0.361 (1.8) |
| 慣性項 | 5.59 (3.6) | | 1.64 (3.7) |
| θ | | 3.26 (4.6) | 2.96 (5.0) |
| スケールパラメータ | 0.268 (3.6) | 0.729 (4.2) | 0.703 (4.4) |
| N | 1739 | 228 | 228 |
| L(0) | -1205.38 | -1205.38 | -1205.38 |
| L(β) | -774.25 | -790.91 | -774.56 |
| ρ ² | 0.358 | 0.344 | 0.357 |
| ρ̄ ² | 0.349 | 0.336 | 0.348 |

b) RP/SP同時推定

従来型のRP/SP同時推定モデルの結果が表-4の右側2列である。両モデルの共通の説明変数として、一人当たりの交通費用、車内旅行時間、車外旅行時間、鉄道の乗り換え回数、業務目的ダミー（鉄道固有）、女性ダミー（鉄道固有）、SPモデルにはRPデータの選択結果を慣性項として用いた。定数項はそれぞれのモデルごとに推定した。

慣性項を除いて同時推定した場合には、乗り換え回数の係数は有意な負値をとり、個人属性についてもRPモデルと同じ符号となったが、定数項は依然異なった符号である。また、誤差項とSP質問で与えられた車内旅行時間が相関を持ったために、車内旅行時間の係数が正值になるような大きなバイアスが生じている。これは慣性項により状態依存を考慮することで解決されている。しかし、この場合にも定数項の値は符号は異なっており、また乗り換え回数の係数も有意でない。

c) 系列相関を考慮したRP/SP同時推定

b) で示した RP/SPモデルでは、慣性項が入った場合「系列相関」と「状態依存」のために理論上パラメータ推定値にバイアスが生じていると考えられる。そこで、系列相関を考慮したRP/SPモデルを推定した(表-5)。

前述の車内旅行時間の係数の大きなバイアスは、系列相関だけを考慮した場合でも改善されており、慣性項を入れることによって大きく改善された。定数項や慣性項の係数の絶対値が、系列相関を考慮することによってかなり小さくなっており、その他の説明変数の説明力が大きくなったモデルと言えよう。また、スケールパラメータも1.0に近くなっており、RPモデルとSPモデルの差が小さくなっていることもわかる。慣性項が無いモデルの場合、系列相関を考慮することによって適合度が大きく上がっていることも注目に値する。

6. おわりに

本稿ではこれまでに提案した RP データと SP データを同時に用いる離散型選択モデルをさらに発展させ、両データの補完的な長所を生かすとともに、状態依存だけでなく既存の方法では考慮できなかった系列相関を明示的に考慮した RP/SP 同時推定モデルを提案した。

2つのケーススタディを行なったが、いずれの場合にも SP データだけを用いたモデルと系列相関や状態依存を考慮しない RP/SP 同時推定モデルでは、パラメータ推定値にバイアスの存在が確認された。これは、SP データには強い状態依存性が存在し、SP データ単独で分析を行なう際にも、この状態依存性を考慮しなければ分析結果に重大なバイアスが生じる危険性が大きいことを示唆している。本研究で行なったケーススタディのように、「新型特急ダミー」や「車内旅行時間」のような政策分析上最も重要な変数の係数が、直感と合わない符号を取ることもある。

RP/SP 同時推定分析やパネルデータを用いたダイナミックな非集計分析では、この問題は「系列相関」と「状態依存」という形で現われ、パラメータ推定値にバイアスを生じさせる。RP/SP 同時推定モデルでは、このバイアスは従来の研究で著者らが行ってきたように RP データの選択結果を慣性項として SP モデルの効用関数に加えること、または個人・選択肢で共通なシステマティックな誤差の相関を明示的に考慮することで解消できることを示した。本論文で提案したような、真の状態依存性を表す慣性項と系列相関を考慮したモデルが、理論上最も望ましい統計的性質を持ったパラメータ推定値を与えるが、今回行なった実証的研究を見るかぎり、慣性項だけを入れて両データを独立と扱う RP/SP 同時推定モデルでも、重大なバイアスの改善という意味では、十分実務の利用に足ると考えられる。

今後、実証的分析を重ねて、本方法論の有効性をより厳密に確認することが重要であろう。

参考文献

- 1) 森川高行：ステイティッド・プリファレンス・データの交通需要予測モデルへの適用に関する整理と展望，土木学会論文集，No.143，pp.4-12，1990.
- 2) 鈴木聡，原田昇：パソコンベースの応答型意識調査手法に関する研究—通勤・通学の鉄道経路選択を対象として，土木計画学研究・論文集，No.6，pp.12-25，1988.
- 3) 藤原章正，杉恵頼寧，中山恵介：パソコンを利用した選好意識調査の一手法の開発，土木学会年次学術講演概要集，No.44/IV，pp.452-453，1989.
- 4) Bradley, M., T. Grosvenor & A. Bouma: An Application of Computer-Based Stated Preferences to Study Mode-Switching in the Netherlands, working paper, 1988.
- 5) Ben-Akiva, M. & T. Morikawa: Estimation of Switching Models from Revealed Preferences and Stated Intentions, *Transp. Res. A*, Vol.24A, No.6, pp.485-495, 1990.
- 6) Ben-Akiva, M. & T. Morikawa: Estimation of Travel Demand Models from Multiple Data Sources, *Transportation and Traffic Theory*, M. Koshi, ed., pp.461-476, 1990.
- 7) Morikawa, T., M. Ben-Akiva & D. McFadden: Incorporating Psychometric Data in Econometric Travel Demand Models, prepared for Banff Invitational Symp. on Consumer Decision Making and Choice Behavior, 1990.
- 8) Morikawa, T., M. Ben-Akiva & K. Yamada: Forecasting Intercity Rail Ridership Using Revealed Preference and Stated Preference Data, *Transp. Res. Rec.*, No.1328, pp.30-35, 1991.
- 9) Amemiya, T.: *Advanced Econometrics*, Harvard University Press, 1985.
- 10) Heckman, J.J.: Statistical Models for Discrete Panel Data, in *Structural Analysis of Discrete Data with Econometric Applications*, Manski & McFadden eds., MIT Press, Cambridge, Massachusetts, pp.114-175, 1981.
- 11) Pas, E.I. & F.S. Koppelman: An Examination of the Determinants of Day-today Variability in Individuals' Urban Travel Behavior, *Transp.*, Vol.14, pp.3-20, 1987.
- 12) Kitamura, R.: Panel Analysis in Transportation Planning: An Overview, *Transp. Res. A*, Vol.24A, No.6, pp.401-415, 1990.
- 13) 森川高行，M. Ben-Akiva：RP データと SP データを同時に用いた非集計行動モデルの推定法，交通工学，Vol.27，No.13，pp.21-30，1992.
- 14) (社)システム科学研究所：在来線高速化需要予測調査報告書，1990.

(1992. 10. 7 受付)

ESTIMATION OF DISCRETE CHOICE MODELS WITH SERIALLY CORRELATED RP AND SP DATA

Takayuki MORIKAWA and Kikuko YAMADA

Disaggregate behavioral travel demand models are usually estimated from revealed preference (RP) or stated preference (SP) data. RP and SP data have complementary characteristics: RP data are more reliable and SP data are more flexible. The authors have developed combined estimation methods for discrete choice models from RP and SP data to enhance advantages of the two types of data. This paper develops the RP/SP combined estimation methodology that explicitly considers state dependence and serial correlation between RP and SP. Two case studies of mode choice analysis shows effectiveness of the proposed methods.