

パネルサンプルの初期摩耗を考慮した動的な買物目的地選択モデル*

Dynamic Choice Analysis Revising Pre-Attrition Biases of Panel Sample

佐々木邦明**, 森川高行***, 杉山幸司****

By Kuniaki SASAKI, Takayuki MORIKAWA and Kohji SUGIYAMA

1. はじめに

個人の行動履歴を表す、パネルデータの非集計行動モデルへの適用に関する研究が近年数多くなされ、パネルデータをモデル構築に用いることの利点と共に、欠点も明らかになってきた。問題としてあげられているものの一つが、サンプルが調査回数を増すにつれ減少してゆくアトリションである¹⁾。一般にパネルサンプルは、転居、死亡などの物理的要因や、調査拒否などによって調査回数が増すにつれて減少していく。このとき、パネル調査より脱落したサンプルが、分析対象となる選択行動及びその説明変数に対し、ランダムでなく、何らかの相関を持っていた場合、パネルサンプルはwaveが進むに連れて分析の対象となる母集団を正確に代表しなくなる。そのようなサンプルから得られたデータをを用いてモデルを推定した場合、そのパラメータは不偏性を失ってしまう。つまり、ランダムサンプリングにより最初のサンプルを抽出したとしても、waveが進むにつれてパネル調査に残りやすい傾向を持ったサンプルが多くなっていき、パネルに残留したサンプルの選択シェアや説明変数の出現確率が母集団と一致しない場合、推定されたパラメータにはバイアスが存在するということである。特に選択結果のシェアの不一致は、選択肢固有定数にバイアスを生じることが知られている²⁾。通常これらのバイアスはアトリ

ションバイアスと呼ばれ、パネル分析の初期より指摘されてきている¹⁾。これまで、このバイアスを考慮した様々な分析手法が開発されており、それらは、各サンプルの母集団中の比率とサンプル全体中の比率を各サンプルの重みとする考え方に従うもの³⁾が多い。

また近年、調査の規模にかかわらずその人的、金銭的面で容易さなどの理由によって、訪問回収にかわり郵送による調査票の回収が用いられることが多くなっている。一般に郵送による回収は、訪問回収と比較して回収率が大幅に低い。郵送回収による調査を分析した研究⁴⁾によると、回収されたサンプルが、既に母集団を表していないと考えられる結果も報告されている。しかし、調査に協力してくれないサンプルは、無視できるほどの大きさではない場合があるにもかかわらず、分析するための情報が全く入手できないときには、無視されることが多かった。

本研究では、郵送回収を行ったときの初期の返答拒否を初期摩耗と定義し、初期摩耗と先に述べたパネル調査からのサンプルの脱落に着目して、その構造をモデル化し、パラメータのバイアスを修正した非集計行動モデルの推定手法の提案を行うことを目的とする。

2. パネルアトリション

(1) パネルアトリションと母集団代表性

非集計行動モデルの推定に用いられる、代表的なサンプリング手法としては、無作為抽出、層別抽出法、選択肢別抽出法があり、それぞれの抽出方法の特性に応じたモデルの推定方法や、パラメータの修正方法が提案されている。それらのほとんどは、各個人の母集団のシェアとサンプル中のシェアの違いに着目したものであり、母集団のシェアが不明であるときには、サンプルを無作為抽出とみなしてモデルを推定する場合が多い。しかし、一般に交通機関選択などの場合は母

* Key Words : 交通行動分析, 調査論

**正会員 修士(工学) 名古屋大学助手工学部土木工学科
(〒464-01 名古屋市千種区不老町)
Tel. 052-789-3565 Fax. 052-789-3738)

***正会員 Ph.D. 名古屋大学助教授 工学部土木工学科
(〒464-01 名古屋市千種区不老町)
Tel. 052-789-3564 Fax. 052-789-3738)

****正会員 修士(工学) 建設技術研究所 東京支社道路本部
道路計画課
(〒103 東京都中央区日本橋本町4-9-11)
Tel. 03-3668-0451 Fax. 03-5695-1883)

集団中の交通機関別のシェアなどが容易に判明するために、先に述べた母集団中のシェアを考慮した推定値の修正が行えるが、非集計行動モデルの適用が可能であると思われる選択行動の一部は、母集団シェアを求めることが困難であったり、不可能な場合も存在する。ここで、パネルサンプルのアトリションについて再び考えてみる。サンプルの脱落が属性や選択に対してランダムな場合には、無作為抽出標本によって推定されたモデルはアトリションの影響を受けず、最尤推定法によって推定されたパラメータは、漸近的有効性と一貫性を保つために、問題は生じない。しかし、これまでのパネルサンプルのアトリションについての研究によると、モビリティの高い人が脱落し易いなど、ある行動特性を持ったサンプルが調査から離脱してゆく、という報告⁵⁾がなされている。つまりパネルサンプルのアトリションの影響とは、ある行動特性を持った個人のサンプル中でのシェアが、母集団でのシェアに一致しなくなるということであり、無作為抽出がある種の層別抽出標本に変化してゆく現象であると考えることができる。このことを考慮して、初期摩耗現象を含めたサンプルの減少をモデル化することによって、これまで無視されることが多かった、無回答者を含めたサンプルの母集団代表性を考慮したモデルの推定が可能になると考えられる。

(2) パネルアトリションによるサンプルの変化

山梨大学の西井研究室が中心となって行っている、甲府パネル調査⁶⁾によって得られたデータを用いて、調査の進行に伴うパネルサンプルの構成の変化を分析する。本調査は郊外立地型の大型ショッピングコンプレックス（以下SC）への来訪者に対して配布され、郵送によって回収されているもので、1989年に第1回の調査が始まり、昨年七回目の調査が行われた。本研究ではその中のwave 3 から wave 5 までのデータを用いて分析を行った。

まずはじめに、アトリションによって各サンプル中の属性の割合が実際に変化しているかを確認した。その中の一例としてwave 3 に初めて調査に参加したサンプルを対象にしてwave 3、4、5での性別構成を図-1に示す。図から分かるようにサンプルの減少率は男性の方が高く女性の方が低い。つまりパネル調査への残留という事象は、性別という属性に対してランダム

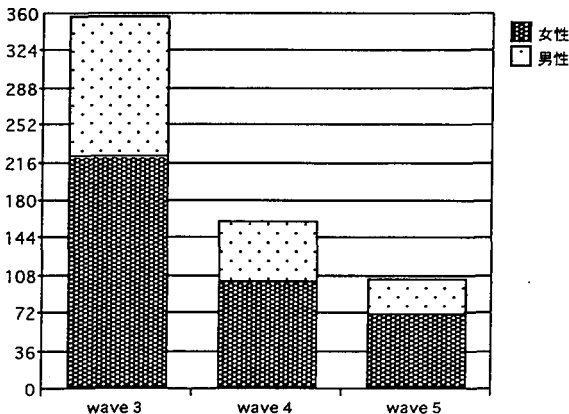


図-1 各waveでのパネルサンプルの性別構成

表-1 グループ別および全サンプルでの選択シェア

wave 3	女性	男性	全サンプル
S C	63.5%	64.9%	64.0%
中心部商店街	8.2%	3.8%	6.5%
最寄りスーパー	28.3%	31.3%	29.5%
計	100%	100%	100%

wave 4	女性	男性	全サンプル
S C	55.3%	54.4%	55.0%
中心部商店街	14.6%	10.5%	13.1%
最寄りスーパー	30.1%	35.1%	31.9%
計	100%	100%	100%

wave 5	女性	男性	全サンプル
S C	49.3%	69.7%	55.8%
中心部商店街	7.0%	6.1%	6.7%
最寄りスーパー	43.7%	24.2%	37.5%
計	100%	100%	100%

に生起していないと考えられる。また調査項目の中の「ここ1～2カ月最もよく利用する買物場所」に対する回答のシェアを、性別ごとと全サンプルで表-1にまとめた。この表をによると、女性と男性ではそれぞれの買物場所の選択シェアに差が存在する。よって性別は選択結果とパネル残留それぞれに何らかの関連があると考えられ、アトリションによるサンプルのバイアスが存在すると考えられる。

3. サンプルのアトリションを考慮した選択モデル

(1) パネル滞留モデル

既存の研究によると、パネルからの離脱現象は、個人がそれまで何回調査に参加しているかによって異なり、パネル調査の回数が増えるにつれて離脱者の割合は減ってゆく、と報告されている²⁾。そこでパネル調査から離脱する行動を以下のように仮定する。個人がはじめて調査票を受け取った時からパネル調査より離脱するまでの時間は、一人一人異なり、かつ同一の分布に従っていると仮定する。従って M を確率変数と考え、その密度関数を $f(\tau)$ 、分布関数を $F(\tau)$ とする。ある個人が調査票を受け取ってから τ 時間経過したという条件のもとで、それに続く微小時間 $\Delta \tau$ 内にパネル調査より脱落する確率は、 $\Delta \tau$ が十分小さいとき、次式で与えられる。

$$P(\tau < M < \tau + \Delta \tau | M > \tau) = \frac{F(\tau + \Delta \tau) - F(\tau)}{1 - F(\tau)} \approx \frac{f(\tau)}{1 - F(\tau)} \Delta \tau \quad (1)$$

ここで、

$$h(\tau) = \frac{f(\tau)}{1 - F(\tau)} \quad (2)$$

とおくと、 $h(\tau)$ は、 τ 時点でパネル調査に滞留しているという条件のもとで、続いておこる微小時間 $\Delta \tau$ にパネル調査から離脱する確率になる。そこで、既存の研究に基づき、 $h(\tau)$ を時間の経過につれて調査から脱落する確率が減少する、つまり $h(\tau)$ を τ の減少関数

$$h(\tau) = a \cdot b^{-a\tau} \quad a, b > 0, \tau > 0 \quad (3)$$

ただし、

a, b : $h(\tau)$ の分布形を定める定数

とおくと、 τ 時点でパネルサンプルが調査から離脱する確率の分布関数は

$$F(\tau) = 1 - \exp\left(\frac{b^{-a\tau}}{\ln b} - \frac{1}{\ln b}\right) \quad (4)$$

で与えられる。これはパネル調査からの離脱を連続時間で表す分布関数であり、パネル調査は時間に対し離散的であるため、離散時点 $t-1$ までパネルに滞留し t 時点で調査より離脱する確率は $F(t) - F(t-1)$ となる。

また、個人 n がパネル調査に滞留する確率は、個人や家庭の属性に影響を受けていることが、同様に報告されている³⁾。これを、(4)式の分布形を定める定数 b が各サンプルの属性によって異なることを意味しているものと考え、(3)式の b に、属性ベクトル X_n を持つ個人 n の属性ベクトルの重み付きの線形和 A_n を、非負条件を満たすようにロジット変換したものを代入すると、(4)式は以下のような形に表される。

$$F_n(t) = 1 - \exp\left[\frac{\left(\frac{\exp(\Gamma X_n)}{\exp(\Gamma X_n) + 1}\right)^{-a}}{\ln\left(\frac{\exp(\Gamma X_n)}{\exp(\Gamma X_n) + 1}\right)} - \frac{1}{\ln\left(\frac{\exp(\Gamma X_n)}{\exp(\Gamma X_n) + 1}\right)}\right] \quad (5)$$

ただし、

X_n : 個人 n の個人属性

Γ : 未知パラメータベクトル

よって個人 n がある時点 t までパネルに滞留し、かつ次のパネル調査に不参加となる確率は

$$P_{n,t,t+1}^*(dropout) = F_n(t+1) - F_n(t) \quad (6)$$

で与えられる。

(2) 系列相関を考慮した選択モデル

各 t 時点での個人 n の選択肢 i に対する効用を以下のよう

$$U_{n,i}^t = BZ_{n,i}^t + \varepsilon_{n,i}^t \quad (7)$$

ただし、

$U_{n,i}^t$: 選択肢 i に対する個人 n の t 時点での効用

$Z_{n,i}^t$: 効用に影響を与える選択肢 i に対する個人 n の t 時点での属性ベクトル

B : 未知パラメータベクトル

$\varepsilon_{n,i}^t$: 個人 n の選択肢 i の効用に対する誤差項

パネル調査は同一個人に対して繰り返し尋ねたデータであり、同一個人から時系列的に得たデータには系列相関と状態依存性が存在することが既存の研究により指摘され⁷⁾、独立で同一の誤差項の仮定は現実的でない。そこで著者ら⁸⁾によって提案された、同一個人に対し時点間で共通の誤差項を系列相関項として導入するモデルを採用して、系列相関の影響を取り除く。また、パネル残留と対象とする選択にも系統的な相関が存在すると考えられる³⁾ため、同一時点でのパネル残留傾向を表す変数と選択肢に対する効用との間に存

在する系列相関、多時点間の選択肢に対する効用の誤差項間に存在する系列相関の2つを仮定し、それらは互いに独立であるとする、それぞれ以下のように定式化される。

$$A_n = \Gamma X_n + \lambda_n \quad (8)$$

$$U'_{n,i} = BZ'_n + \lambda_n + \theta_{n,i} + \varepsilon'_{n,i} \quad (9)$$

ただし

λ_n : A_n と各時点の効用の間に存在する系列相関

$\theta_{n,i}$: 各時点の効用間に存在する系列相関

つまり、 λ は $U'_{n,i}$ と A_n に共通するの観測不能な要因

などを表し、 θ は同様の原因によって選択モデルの誤差項の時系列的な相関を示すものである。また、 λ は時間的に安定していると仮定することで、多時点間のパネル残留傾向に θ は含まれない。これらより、2つの系列相関項が既知のもとでの個人 n が t 時点で選択肢 i を選ぶ確率は、(9)式の $\varepsilon'_{n,i}$ にI.I.D.ガンベル分布を仮定することによって、選択確率は多項ロジット型で表され、以下ようになる。

$$P_{n,i}(i|\lambda, \theta) = \frac{\exp(BZ'_{n,i} + \lambda_n + \theta_{n,i})}{\sum_{k \in C_n} \exp(BZ'_{n,k} + \lambda_n + \theta_{n,i})} \quad (10)$$

ただし、 C_n は個人 n の選択肢集合

パネル滞留モデルについても λ が既知のもとでの個人 n が $t+1$ 時点でパネルから脱落する確率は

$$P^*_{n,t,t+1}(\text{dropout}) = F_n(t+1|\lambda_n) - F_n(t|\lambda_n) \quad (11)$$

と表される。

(3) パネルアトリションを考慮した選択モデル

背景でも述べたように、パネルアトリションの影響によって、パネルサンプルが調査に残りやすい層が多く抽出された層別抽出に近似できるとすると、各時点でのサンプルのアトリションを考慮した尤度は、層別抽出標本によるモデルの推定と同様の考え方によって、摩擦傾向の同一な個人の属する抽出層をグループと定義すると以下の式で表すことができる2)。

$$L(\beta) = \prod_{g=1}^G \prod_{n=1}^{N_g} \prod_{i \in C_n} \frac{P(i|Z_n, \beta)^{\delta_{in}} P(Z_n) H(g)}{Q(g)} \quad (12)$$

ただし、

β : 未知パラメータ

$P(i|Z_n, \beta)$: β および Z_n が与えられたときの選択肢 i の選択確率

δ_{in} : 個人 n が選択肢 i を選択した場合1、それ以外0
 $H(g)$: グループ g のサンプル中でのシェア
 $Q(g)$: グループ g の母集団中でのシェア
 N_g : グループ g の中のサンプル数
 G : グループ数

ここで、パネル滞留モデルより、摩擦傾向が同一な個人が属するグループの t 時点までのパネル滞留確率

$$P^*_{t,g}(\text{stay}) = 1 - F_g(t) \quad (13)$$

が与えられると、各 t 時点で以下のような関係が成立する。

$$\hat{N}_{0,g} = N_{t,g} / P^*_{t,g}(\text{stay}) \quad (14)$$

ただし

$\hat{N}_{0,g}$: 調査票配布時点でのグループ g の人数の推計量

$N_{t,g}$: t 時点でのグループ g の残留人数

(14)式より、グループ g の母集団内でのシェアの推計値

$Q(g)$ が、以下のように計算できる。

$$Q(g) = \hat{N}_{0,g} / \sum_{j=1}^G \hat{N}_{0,j} \quad (15)$$

ただし、この $\hat{N}_{0,g}$ はどの時点をとるかによって変動すると考えられるため、本来は各時点での推計値の平均値をとることが望ましいと考えられるが、今回提案したモデルでは簡単のため、各時点での推計値をそのまま各時点でのシェア計算に用いた。

以上より、系列相関及びパネルアトリションを考慮した、グループ g に属する個人 n のパネル調査への残留、脱落および各時点の選択の同時確率は、

$$P_{g,n} = \int \int \prod_{t=1}^T \left(L_n^{-1} \cdot C'_n \right) \cdot \left(P^*_{g,n,T,T+1}(\text{dropout}|\lambda_n) \right) \cdot f(\lambda) \cdot f(\theta) d\lambda d\theta \quad (16)$$

で表される。ただし

$$L'_n = \frac{P^*_{g,n,t,t+1}(\text{stay}|X_n, \lambda_n) P(X_n) H(g)}{Q(g)}$$

$$C'_n = \frac{P_{g,n,t}(i|Z_n, \lambda_n, \theta_n)^{\delta_{in}} P(Z_n) H(g)}{Q(g)}$$

T : 個人 n が残留した最大wave

$f(\lambda)$: λ の同時確率密度関数

$f(\theta)$: θ の確率密度関数

ただし、(16)式中 $t=1$ の場合の L_n^0 とは、仮定のwave0での人数が定式化できるため、初期摩擦によりパネル調査より脱落したサンプルをモデル化したものである。最終的には(16)式で表される個人の選択確率を、グループ内および各グループについて積をとったものを尤度

関数として最尤推定法を用いて全ての未知パラメータを同時に推定する。

4. 事例研究

3. で提案したモデルの有効性を確認するために、
2. で用いた甲府パネル調査データを用いて、事例研究を行う。甲府パネル調査データには、以下のような項目が含まれている。

- 1)最近1～2カ月の休日に最もよく利用する買物場所
 - 2)回答者の属性や世帯の属性など
 - 3)自宅から買物場所への距離や満足度など
- これらのデータより、買物目的地選択行動を分析対象とする。買物目的地選択モデルはSC、甲府市中心部の商店街、最寄りのスーパーの3項選択とし、その説明変数は同じデータを用いた著者らの既存の研究⁹⁾と同様のものを用いる。摩擦モデルは、定数項の他に2. でサンプル中の選択シェアの変動に特徴が見られた性別をダミー変数として説明変数に用いた。この摩擦モデルの説明変数については、パネル滞留行動に関連する属性を全て含むように設定されねばならないが、抽出層の数が、考慮する属性の数とその属性のカテゴリー数に応じて指数的に増えるため、今回の事例研究では、提案モデルの有効性を確認することを目的として、2. で用いた性別を例にして検討を行った。

パネルアトリションと系列相関を考慮したモデルの推定結果を表一2に示した。それぞれ変数名中の数字はwave数を表している。また比較のため、系列相関だけを考慮し、パネルアトリションを考慮しないモデルの推定結果を表一3に示した。ただし、アトリションを考慮したモデルとの比較検討のために、性別を説明変数として加えたモデルと、加えないモデルの2つをそれぞれ表中に示した。これらの推定結果を比較すると、性別を説明変数として加えることによりそれほど定数項の値は変化しないが、アトリションを考慮することで定数項の値に大きな違いが現れ、パネルアトリションを考慮したモデルは、wave4の定数項が有意ではなくなった。同様にスケールパラメータはアトリションの影響を考慮すると、waveが進むごとに誤差項の分散が大きくなるように変化した。これらはともに女性のサンプルの影響が相対的に小さくなったことによるものと考えられる。また、その他の説明変数のパ

表一2 パネルアトリションを考慮した買物目的地選択モデルの推定結果

変数名	推定値	t 値
SC定数項 3	0.696	2.3
SC定数項 4	0.168	0.5
SC定数項 5	-0.0580	-0.4
距離	-1.05	-8.5
交通利便性満足度	0.321	4.5
買物利便性満足度	0.251	2.6
スケールパラメータ 3 4	0.739	3.7
スケールパラメータ 3 5	0.425	3.6
アトリション定数	-1.68	-19.1
性別ダミー	1.07	9.9
a (定数)	0.508	108.0
その他の統計量 観測数620 p ² : 0.150		
ただし 性別ダミー:	女性 1, 男性 0	

表一3 系列相関を考慮した買物目的地選択モデルの推定結果 (性別なし、あり)

変数名	推定値 (t 値)	推定値 (t 値)
SC定数項 3	1.11 (5.0)	1.13 (4.5)
SC定数項 4	0.762 (2.4)	0.879 (2.5)
SC定数項 5	0.296 (0.8)	0.316 (0.9)
性別		-0.556 (-2.6)
距離	-0.626 (-5.7)	-0.0470 (-0.7)
交通利便性満足度	0.154 (2.8)	0.253 (4.2)
買物利便性満足度	0.280 (3.9)	0.141 (1.9)
スケールパラメータ34	1.13 (3.3)	1.61 (2.3)
スケールパラメータ45	0.798 (3.2)	0.767 (2.1)
その他の統計量	p ² 0.269	p ² 0.391
サンプル数 620		
ただし 性別ダミー:	女性 1, 男性 0	

ラメータについては、交通利便性満足度と買物利便性満足度の推定値の大小が逆転していることがあげられるが、これは性別を効用関数の説明変数に加えることによって生じることがわかる。また、距離のパラメータの相対的の重みに変化が生じたことは、性別を効用関

数の説明変数に用いた場合とは異なる傾向を示した。つまり、性別を効用関数の説明変数として用いると、説明変数との相関によるバイアスの影響が推定値に現れるが、アトリションバイアスを考慮したモデルは、性別を効用関数の説明変数に用いたモデルと比較して、他の説明変数の推定値だけでなく選択肢固有定数にも影響を与えることがわかる。またパネル残留モデルのパラメータは女性ダミーの符号が正になっていることから、女性がパネルに残留しやすいことを示しており、予想されたとおりの結果である。適合度指標 $\bar{\rho}^2$ はアトリションを考慮したモデルの方が大幅に低いのは、パネル残留モデルも内包し、かつ仮定のwave0からの滞留も表現していることに起因すると思われる。

5 まとめ

本研究はアトリション現象をモデル化することで、パネルサンプルのアトリションによるバイアスを修正したモデルの提案を行った。そのとき、最初から調査への参加拒否したサンプルの行動を初期摩擦と定義し、その影響を考慮したパネルサンプルの重み付を行った。より具体的には、サンプルのパネル調査への残留傾向に影響する属性の線形和を用いてパネル滞留モデルを構築し、それより母集団での属性シェアを定式化し、層別抽出によるサンプルの重みを考慮した尤度関数に適用して、アトリションの影響を考慮したパネル滞留モデルと選択モデルのシステムを構築した。そのとき、パネル残留と時系列的選択に個人特有の誤差項を導入し、系列相関の影響をパラメータの推定量から除去した。

買物目的地選択データを用いた事例研究により、パネルアトリションを考慮しないモデルと比較して、各waveの定数項、各waveの誤差項の分散、説明変数などの推定値が変化した。同時にパネルアトリションに影響する属性を効用関数に代入したモデルとの比較では、パネルアトリションを考慮したモデルと同様の傾向を示す推定値もあったが、アトリションを考慮することで、効用関数に属性を導入するモデルではほとんど変化のなかった定数項の推定値と誤差分散の比が大幅に変化した。このことから、提案したモデルは、説明変数を追加するモデルが取り除くバイアスだけでなく、それだけでは除去できないバイアスも取り除くと考え

られる。ただし、本研究はただか3回のパネル調査のデータを用いてモデルを推定したものであって、パネルの滞留現象はこの先に一層変化があることが予見され、パネル滞留モデルの信頼性はその説明変数の少なさとともに未だ高いとは言えない。また、今後、より多くの事例研究による検証が必要であると考えられる。

参考文献

- 1) R. Kitamura: Panel Analysis in Transportation Planning - An Overview -, Transportation Research-A, vol.24-A, No.6, pp.401-412, 1990.
- 2) 土木学会編：非集計行動モデルの理論と実際，1995.
- 3) 西井和夫ら：パネルアトリションを考慮した買物場所選択モデル：甲府買物パネルデータを用いて，土木計画学研究・論文集，No.12，pp.389-396，1995.
- 4) 森川高行，田中小百合：トリップ分析による郵送方式パーソントリップ調査の適用性に関する研究，土木計画学研究・講演集，No.16(1)，pp.305-310，1993.
- 5) J.M. Golob and L. J. M. Schreurs : The Design and Policy Applications of a Panel for Studying Changes in Mobility Over Time, Behavioural Research for Transportation Policy, pp.81-95, 1986.
- 6) 西井和夫，岩本哲也：ショッピングコンプレックス来訪者の買物行動特性の基礎的分析，土木計画学研究・講演集，No.13，pp.975-982，1990.
- 7) Daganzo, C.F and Y. Shefi: Multinomial Probit with Time-Series Data : Unifying State Dependence and Serial Correlation Models, Environment and Planning A, Vol.14, pp.1377-1388, 1982.
- 8) 森川高行，山田菊子：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，土木学会論文集，No.476/IV-21，pp.11-18，1993.
- 9) 佐々木邦明，森川高行，杉本直：潜在セグメントを考慮した動的な休日買物目的地選択分析，土木計画学研究・論文集，No.12，pp.397-404，1995.

パネルサンプルの初期摩耗を考慮した動的な買物目的地選択モデル

佐々木邦明・森川高行・杉山幸司

パネルデータを用いる際の問題点の一つとして、調査回数が増えるに従ってパネルサンプルが減少していくアトリションがあげられる。一般に減少するサンプルの属性と選択結果の間には相関があり、そのようなサンプルを用いてモデルを推定すると、パラメータに偏りが生じてしまう。そこで本研究では、サンプルの減少状況を記述する磨耗モデルを用いて各属性ベクトルの母集団中でのシェアを定式化し、これを層別抽出の尤度関数に適用することで目的地選択モデルを同時に推定し、アトリションバイアスの影響を取り除くモデルシステムを構築した。このシステム中では、パネル調査に最初から協力しなかったサンプルの存在を初期摩耗と定義し、それも考慮にいれて母集団中のシェアを計算するようなシステムになっている。そして、提案したモデルの事例研究を甲府市で行われた買物目的地選択調査のデータを用いて行い、その有効性を検証した。

Dynamic Choice Analysis Revising Pre-Attrition Biases of Panel Sample

Kuniaki SASAKI, Takayuki MORIKAWA and Kohji SUGIYAMA

When modeling travel behavior using panel data, the most serious problem is sample attrition. Unless the sample attrition take place randomly to the travel behavior, that cause biases on the estimated parameters, named attrition bias. The purpose of this study is to consist a dynamic model system considering sample attrition using a panel data. The model system we propose consists of two parts, attrition model and choice model. Each model describes attrition and choice behavior, respectively. Then we applied the model system to the shopping destination choice data on non-workday to confirm the efficiency of the model system.
