

行動論的背景を持たせた生存時間モデルによるパネルアトリションバイアスを修正した選択モデル*

Discrete Choice Model considering Panel Attrition Biases using a Duration Model based on the Behavior Theory*

佐々木邦明**

Kuniaki SASAKI**

1. はじめに

交通需要予測が必要とされる局面がハードウェア整備型からソフト的政策分析へと変化するにつれて、動的特性を考慮した交通行動分析の必要性が高まり、パネルデータを用いた交通行動の動的分析が、研究レベルでは盛んに行われている。パネルデータは、その情報量の多さによるさまざまな長所を備えている反面、パネル調査特有の問題もいくつか存在する。そのうちの一つに、調査を繰り返すにつれてサンプルが減ってゆくことに起因するアトリションバイアスの問題がある。

アトリションバイアスについては、その問題存在の明確さより、これまでさまざまな研究が行われてきている。アトリションバイアスを修正した交通行動分析手法の主なものには、母集団での各属性の分布とサンプルでの属性分布の違いを用いて、サンプルに重みを付けるタイプと、選択行動とパネル残留行動相互に関連を持たせるタイプの2種類が存在する。著者らは先に、選択モデルの効用関数とパネル残留を規定する変数に相関を持たせ、生存時間モデルを用いて、観測が困難な真の母集団での属性分布を推定したうえで、各サンプルに重みをつけるシステムを提案した¹⁾。このシステムでは、選択モデルは通常のランダム効用理論から導出されるロジットモデルを用い、パネル滞留行動は、行動理論的背景を持たない生存時間モデルによって規定していたため、行動分析モデルとしての一貫性に欠けていた。そこで、本研究ではパネル残留モデルを、効用理論に基づいて導出することで、モデルシステム全体としての一貫性を持たせることを試みる。

* Key Words 交通行動分析, 調査論

** 正会員, 修士(工学), 名古屋大学大学院工学研究科

〒464-01 名古屋市中千種区不老町

Tel 052-789-3565 Fax 052-789-3738

2. 従来の研究

信頼性分析で用いられている生存時間モデル(Duration Model)は、機械などの寿命時間の分析を目的として開発され、計量生物学や経済学、マーケティングなどの分野で用いられてきた。近年、交通行動の分野でも、活動継続時間の分析や活動の時間配分などの分析に用いられる例が見られるようになってきた。そのレビューとしてはHensher and Mannering²⁾が挙げられる。最近の研究では、張ら³⁾はパネル調査期間の設定の基礎分析として、生存時間モデルを用いて調査参加意志の継続期間を分析している。同じ研究グループは再発可能な事象を取り扱うMultiple-spell Duration Modelを用いてP & R実験への参加を分析している⁴⁾。Bhat⁵⁾はノンパラメトリックにハザードベースが変化する生存時間モデルを用いて、個人の異質性を考慮した帰宅時買物行動の継続時間を分析している。これらはいずれも個人の行動原理に基づいて生存時間モデルを構築したものではなく、状況記述モデルとして生存時間モデルを用いている。小林ら⁶⁾は生存時間モデルにランダム限界効用理論に基づいた行動論的背景を与え、どのような仮定の下でどのタイプの生存時間モデルが導出されるかを分析している。本研究はこの研究を参考に、生存時間モデルに行動論的背景を与える。

3. 生存時間モデル

(1) 生存時間モデルの特徴

生存時間モデルは微小時間に故障が発生する確率を示すハザード関数の形に応じて、D F R (減少)型, C F R (一定)型, I F R (増加)型の3種類に分けられる。D F R型は初期的な不良によつ

て製品寿命が規定される現象を表し、時間とともに故障率が低下して行く。CFR型は故障発生がランダムに生起するもので、故障の発生確率は時間によらず一定である。これは、多くの部品の組み合わせによるシステムの故障発生現象に適用される場合が多い。IFR型は製品が磨耗、腐食、疲労などの原因で故障する現象を記述するのに適しており、時間の経過とともに故障の発生確率が高まって行く。本研究で分析対象とするパネル調査への参加行動は、これまでの事例研究の結果から、パネルからの離脱確率は徐々に減少して行くことが報告されているため、DFR型の生存時間モデルを考える。

(2) パネル滞留時間モデルの導出(その1)

まず、DFR型生存時間モデルの基本的な考え方により、パネル調査に滞留する時間の分布関数を導出する。パネルからの離脱確率は、個人がそれまで何回調査に参加しているかによって変動し、パネル調査への参加回数が増えるにつれて離脱の割合は減ってゆき、最終的には安定した状態に到達して行くと報告されているものが多い(例えば、Wissen and Meurs⁷⁾)。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。個人が調査票を受け取った時点からパネル調査より離脱するまでの時間は、一人一人異なるが、同一の分布に従っているとすると、その密度関数を $f(\tau)$ 、分布関数を $F(\tau)$ とすると、ハザード関数は

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

となる。

ここで $h(\tau)$ を τ の減少関数

$$h(\tau) = p\tau^{p-1} \quad p: 0 < p < 1 \text{ の定数} \quad (2)$$

と仮定すると、(1)式を積分することにより

$$F(\tau) = 1 - \exp(-p\tau) \quad (3)$$

というワイブル分布の特殊形が導かれる。

(3) パネル滞留時間モデルの導出(その2)

(2)でパネル調査からの離脱現象を、機械の故障と同様の原理に従っているとしてみなして、ワイブル分布を導出したが、このパネル滞留時間モデルに行動論的意味あいはない。そこで、行動論的背

景から生存時間モデルを導出する。まず、パネル滞留現象に以下の2点を仮定する。

- ・パネル調査へ参加することに対する各個人のコスト c は時間の増加に対して逓増して行く。
- ・パネル調査に参加することによって得られる限界効用 v は経過した時間に関わらず一定である。

この仮定の下で、パネルへの参加・不参加を決定する行動原理を以下のように仮定する。意思決定者は、パネル調査が行われた各時点で、限界費用と限界効用を比較し、限界効用が限界費用より高い場合にはパネル調査に参加し、限界費用が限界効用を越えていた場合にはこの調査に参加しない。

ここで簡単のため、各意思決定者のパネル調査参加の限界効用は、ある特定の確率分布に従い、限界費用関数は全意思決定者で同一であるとする。これは、パネル調査に参加することによって得られる限界効用が全意思決定者が同一の関数で低減し、パネル調査に参加する限界費用が特定の分布に従うと置き換えることができる。これを図に示すと図-1のようになる。

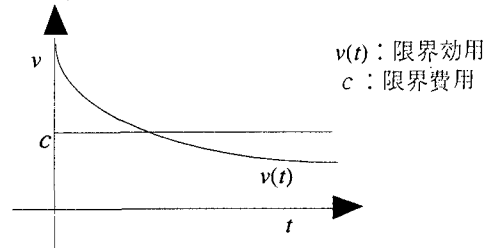


図-1 限界費用曲線と限界効用曲線

以下この仮定に沿ってパネル調査への参加確率を求める。ここで、それぞれの関数型を特定化し、パネル滞留時間分布を求める。パネル調査への参加による限界効用の関数型を

$$v(t) = p \times \exp(-pt) \quad p: 0 < p < 1 \text{ の定数} \quad (4)$$

とする。このとき

$$v(t) = \begin{cases} p & t = 0 \\ 0 & t = \infty \end{cases} \quad (5)$$

となる。

パネル調査への参加の限界費用の分布を、上限を $v(0) = p$ 、下限を0とする一様分布を仮定する。ただし、完全調査不参加者の行動を記述するため、WAVE 1を $t=0$ でなく、 $t=\alpha$ $\alpha > 0$ であるとし、

α 時点で限界効用が限界費用を下回っている場合は調査に最初から参加しないとする。分布の上・下限を特定しことにより、限界費用の分布は

$$f(c) = \frac{1}{p} \quad (6)$$

となる。ここで、ある時点 t においてすでにパネルから脱落している確率は、 $f(c)$ を $v(t)$ に沿って 0 から t まで積分することにより得られる。この積分を評価すると

$$F(t) = 1 - \exp(-pt) \quad (7)$$

となり、先に故障確率に基づいて求めた分布関数と一致する。

(4) ワイブル分布を用いたパネル参加モデル

これまでのパネル残留に関する研究で、パネル残留行動が個人のライフサイクルステージや属性に影響を受けていることが報告されている。そこで、パネル残留に影響を与える変数 A_n を以下のように定める。

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

X_n : 個人 n の属性ベクトル

Γ : 未知パラメータ行列

(7) 式のパネル残留時間の分布関数の p に、変数 A_n をロジット変換して代入し、パネル残留モデルの分布関数を以下のように変形する。

$$F_n(t) = 1 - \exp \left[\left(- \frac{\exp(\Gamma X_n)}{\exp(\Gamma X_n) + 1} \right)^t \right] \quad (9)$$

このモデルは、連続時間上で定義されているが、現実のパネル調査は離散的に行われる。そこで、離散時間である $t-1$ WAVE から t WAVE の間にパネルから脱落する確率はこの分布関数の差である

$$P_{t-1,t}^*(dropout) = F(t) - F(t-1) \quad (10)$$

で与えられる。

(5) パネルアトリションを考慮した選択モデル

(10) 式のパネル残留モデルを用いて、パネルアトリションを考慮した選択モデルを構築する。選択モデルは通常のロジットモデルを用い、同一個人間に誤差項の系列相関を仮定している。アトリションバイアス修正の基本的な考え方は、パネル残留モデルから母集団での属性分布を求め、各サンプルにウェ

イトを付けるというものである。また、そのときに、パネル残留に影響する変数と選択行動を規定する潜在変数(効用)の間に相関関係を仮定し、双方に共通の確率項を導入する。そして、ウェイトが与えられたもとの全 WAVE でのパネル残留・選択同時確率を尤度関数として、未知パラメータを推定する。以下に尤度関数を示す。

$$P_n = \int \prod_{t=1}^T (L_n^{t-1} \cdot C_n^t) \cdot (P_{n,T,T+1}^*(dropout | \lambda_n)) \cdot f(\lambda) d\lambda \quad (11)$$

ただし、

λ_n : 個人 n の効用と A_n に影響を与える確率項

L_n^t : 残留確率 $\frac{P_{n,t,t+1}^*(stay | X_n, \lambda_n) P(X_n) H(X_n)}{Q(X_n)}$

C_n^t : 選択確率 $\frac{P_{n,t}(i | Z_n, \lambda_n) P(Z_n) H(Z_n)}{Q(Z_n)}$

Z_n : 選択に影響を与える変数

T : 個人 n がパネルに滞留した最大 WAVE 数

$f(\lambda)$: λ の確率密度関数

式中に存在する L_n^0 は、調査票配布時つまり母集団での属性分布の推定値を用いて、最初から調査に参加しなかったサンプルの行動を示している。最終的には式を $f(\lambda)$ に多変量独立標準正規確率密度を仮定し数値的に積分を評価することで、個人の選択確率をパラメトリックに定式化し、最尤推定法を用いて各未知パラメータを推定することになる。この推定法による得られるパラメータは一致性を持つ。このモデルについては著者らの研究¹⁾に詳しく述べられている。

4. 事例研究

本章では、甲府パネルデータ⁸⁾を用いて先に述べたモデルの事例研究を行う。甲府パネルデータは、現在まで 8 WAVE のデータが揃っている。データの制約上 WAVE 3 ~ WAVE 6 のデータを用いてモデルを特定化した。買物目的地選択のモデルは著者らの前研究と同じ、対象 SC、中心街、最寄りスーパーの 3 項選択とし、以下の様な 3 種類の説明変数と SC と中心街の定数項を導入した。

1) 自宅から買物場所への距離

2) 各買物目的地に対する交通利便性満足度指標

(10段階評価)

3) 各買物目的地に対する買物利便性満足度指標

(10段階評価)

消耗モデルの説明変数は、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、モデルの構造上説明変数の数に対して計算負荷が指数的に増えるため、今回の事例研究では、サンプル中の選択シェアの変動に特徴が見られた性別と定数項の線形和とした。推定結果は表-1の通りである。

表-1 モデルの推定結果

	推定値	t 値
SC定数項3	0.0239	0.1
SC定数項4	-1.41	-3.4
SC定数項5	-0.254	-0.9
SC定数項6	-0.518	-1.5
中心街定数項3	-0.893	-1.6
中心街定数項4	-3.03	-2.0
中心街定数項5	-0.923	-2.0
中心街定数項6	-2.49	-2.1
距離	-0.531	-4.2
交通利便性満足度	0.105	2.1
買物利便性満足度	0.245	3.3
スケールパラメータ34	1.79	2.8
スケールパラメータ35	0.685	2.0
スケールパラメータ36	1.28	1.9
消耗定数	-0.356	-12.9
性別	-0.00620	-0.3

この推定結果を通常のモデルと比較すると、通常のモデルではSCの定数項が有意なものが多く、中心街定数項が有意でないものが多かったが、今回提案したモデルの推定結果では、ほぼ逆の結果になった。これはサンプルにウェイトを付けることで、選択シェアが変動したことによるものと思われる。また、スケールパラメータの傾向として、誤差分散がWAVE3に対して相対的に小さくなった。特にWAVE6の誤差分散の大小関係が通常のモデルと比較して、逆の結果を示している。距離と2つの満足度変数のパラメータはアトリションによる影響をほとんど影響を受けていないことが読みとれる。

4. おわりに

本研究では、行動論的前提条件より導出した生存時間モデルを用いて、パネルアトリションによるバイアスを修正した選択モデルを提案し、その事例研究を行った。その結果は、事前に予想されたようなパラメータの変動を示し、このモデルがアトリションによるパラメータのバイアスを修正している可能性が示された。ただし、この事例研究では、アトリションを規定する変数が有意でなく、モデルの有効性が低下したと考えられるため、今後より当てはまりのよい消耗モデルの構築が必要であると考えられる。また、本研究で導出した生存時間モデルは数々の仮定を含んでいる。この仮定が行動論的にどのような意味を持っているかを本研究では明確にしていない。よってこれらの仮定の適用可能な条件の検討も今後の課題である。

参考文献

- 1) 佐々木邦明、森川高行、杉山幸司：パネルサンプルの初期磨耗を考慮した動的な買物目的地選択モデル、土木計画学研究・論文集、No.13、pp.595-602、1996。
- 2) Hensher, D. A. and Mannering, F.: Hazard-based duration models and their application to transport analysis, Transport Review, 14, pp. 63-82, 1994.
- 3) 張峻屹、杉恵頼寧、藤原章正、奥村誠：SPパネル調査の調査期間決定のための基礎分析、土木計画学研究講演集、No.19(2)、pp.795-798、1996。
- 4) 重松史生、藤原章正、杉恵頼寧、張峻屹：Duration modelによるパーク・アンド・ライド社会実験への参加行動の分析、土木計画学研究・講演集、No.19(2)、pp.849-852、1996。
- 5) Bhat, C. R.: A Hazard-based Duration Model of Shopping Activity with Nonparametric baseline Specification and Nonparametric Control for Unobserved Heterogeneity, Transportation Research, Vol.30 (B), No.3, pp.189-207.
- 6) 小林潔司、喜多秀行、後藤忠博：ランダム限界効用に基づく滞在時間モデルの導出、土木計画学研究・講演集、No.19(2)、pp.241-244、1996。
- 7) van Wissen, L. J. G. and Meurs, H. J.: The Dutch Mobility Panel: Experiences and Evaluation, Transportation, Vol.16, pp.99-119, 1989.
- 8) 西井和夫、近藤勝直、古屋秀樹、梶木秀典：多時点パネルのアトリションバイアスに関する基礎的考察：甲府買物パネルデータを用いて、土木計画学研究・講演集、No.19(2)、pp.791-794、1996。