

# 多時点パネルのアトリションバイアスとその修正法に関する研究：

## 甲府買物パネルデータを用いて\*

Panel Attrition and the Correcting Method for its Biases :

Using Multi-Wave Kofu Shopper Panels\*

西井和夫 \*\*, 近藤勝直 \*\*\*, 古屋秀樹 \*\*\*\*, 栃木秀典 \*\*\*\*\*

By Kazuo NISHII\*\*, Katsunao KONDO\*\*\*, Hideki FURUYA\*\*\*\* and Hidenori TOCHIGI\*\*\*\*\*

### 1. はじめに

人間の交通行動の時間的変化に対応できるパネル分析は、今後の交通政策による交通需要の予測ならびに評価に大きく寄与すると期待されている。これは、パネルデータに基づいた分析を通じて、個人の交通行動の動的な分析が可能 (例えば、何らかの交通政策がなされた場合には、個人の交通行動の経年的な変化を詳細に記述することができる) であり、また将来予測を行う際は、個人間の異質性やその動的特性に基づく予測モデルの構築により、予測精度の向上が期待できるからである。

このようにパネル分析は、動的な交通行動分析を行おうとするとき、有効な分析ツールとなり得る。しかし、一方で、調査技法としていくつかの課題点があり、調査回数を経るごとにサンプルが減少していくパネルアトリションはその代表的なもののといえる。パネル調査では、時点の進行に伴い、転居や回答拒否などにより被験者が次第に脱落 (drop out) してサンプルが減少 (消耗 : attrition) してしまうという状況が発生する。その結果、サンプルの属性や行動結果の分布に偏り (bias) が生じ、パネルサンプルが母集団を代表していることを単純に期待することができなくなる。これがアトリションの問題であり、パネルデータを用いモデルを推定する場合、この「アトリション (バイアス)」を修正しなければ、母集団とは異なった予測をしてしまう結果となる。筆者らはこれを修正する方法としてWESML法を基礎とした参加確率の逆数で重みづけする方法<sup>1)~3)</sup>

を提案してきた。また、杉恵ら (1993) は参加確率を用いて表現されるバイアス修正項を説明変数の一つとして取り入れたモデル<sup>4)</sup>を提案している。また、佐々木ら (1996, 1995) は母集団のシェアとサンプルシェアの違いを考慮したモデルの開発<sup>5)</sup>や潜在セグメント<sup>6)</sup>を用い個人の嗜好の違いを考慮したモデルの提案を行っている。このようにバイアスを考慮した分析はすでにいくつかの例があるが、多時点にわたるパネル調査への参加・不参加を説明できる適当なモデルやバイアス修正法の開発は十分でない状況といえる<sup>7)</sup>。

そこで本研究では、これまでのアトリションに対する研究事例を踏まえながら、とくに多時点のくり返しデータ上のアトリション構造のモデル化とそれにもとづく具体的なバイアス修正法の適用を通じて、提案手法の有効性を明らかにしていく。

### 2. アトリションの実態把握

一般に、パネルサンプルが調査時点を経て減少していく場合、もしそのサンプルの脱落が属性や選択に対してランダムであれば、モデルのパラメータ推計においてこのアトリションの影響を受けない。しかし、もし脱落するサンプルが分析対象となる選択構造およびその説明変数に対して、何らかの相関を持つならば、パネルサンプルはwaveが進むにつれて分析対象母集団を正確に代表しなくなる。つまり、パネルサンプルのアトリションの影響とは、ある特徴を持ったグループのサンプル中でのシェアが母集団のシェアに一致しなくなることを指し、これは無作為抽出を仮定していた標本抽出がある種の層別抽出法によるサンプリングに変化していく現象とみなすことができる。

ここで、本研究で分析対象とする甲府買物パネルデータの説明をすると、1989年の第1回目の調査

\*Keywords : 交通行動分析, パネル分析, パネルアトリション

\*\* 正会員 工博 山梨大学工学部土木環境工学科  
(甲府市武田 4-3-11 TEL&FAX 0552-20-8533)

\*\*\* 正会員 工博・商博 流通科学大学情報学部  
(神戸市西区学園西町 3-1 TEL078-794-3517 FAX078-794-3054)

\*\*\*\* 正会員 工博 山梨大学工学部土木環境工学科  
\*\*\*\*\* 学生員 山梨大学工学部土木環境工学科

に始まり、昨年1996年の調査までの8年間、合計8時点にわたるデータセットとなっており、この調査は、調査日当日にSC店頭において直接調査票を配布する来訪調査と、そうして得られた来訪サンプルに対してその後の調査時点において調査票を郵送配布するパネル調査の2種類から成り立つ。表-1は、甲府買物パネル調査の具体的なサンプル数とアトリションの割合を示したものである。表中の斜線部は新規調査参加者 (refreshments) であり、時点の経過とともに右方向に進んでいく。

下段はt-1時点 (前時点) のサンプル数に対するt時点での調査への不参加 (drop out) の割合である。このアトリションの割合を詳しく眺めると、来訪時点 (t時点) から次時点 (t+1時点) にかけてのアトリション割合が、それ以降におけるアトリション割合に比べて高いことがわかる。以下では、このような性質をもう少し詳しく眺めてみることにする。

本研究では、パネルデータとしての最初の母集団である来訪時のサンプルをt時点とし、以後時点が進むごとにt+1時点、t+2時点としている。ここで、パネル調査に参加してもらう時、図-1に示すように参加時点 (t時点) から次時点 (t+1時点) に移行する際、調査側から被験者に対して調査票を郵送し、調査に参加してもらうが、このとき郵送できないサンプルがwaveによってはかなり多くを占めている場合がある。郵送不可者は大きく分けて、転居によるものと調査への拒否感からのもの (住所・氏名の無記入による) に分類できる。このとき郵送できないサンプルを数多く含むため、それ以降のアトリションとは性質が異なることが予想される。したがって本研究では、これを初期アトリションと定義し、t+1時点以降のアトリションと区別して扱うことにする。

また、郵送できなかったサンプルについての具体的な集計結果を図-2に示す。waveによって差があり、特にwave7では約30%のサンプルが郵送できていない。これよりwave7における初期アトリション全体の割合は、ここで比較した4waveの中では大きくなっていると考えられる。転居のため

表-1 甲府買物パネル：参加者数及びアトリションの割合

|       | wave1<br>1989 | wave2<br>1990 | wave3<br>1991 | wave4<br>1992 | wave5<br>1993 | wave6<br>1994 | wave7<br>1995 | wave8<br>1996 |
|-------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| wave1 | 653           | 221<br>66.2%  | 153<br>30.8%  | 117<br>23.5%  | 86<br>26.5%   | 71<br>17.4%   | 52<br>26.8%   | 44<br>15.4%   |
| wave2 |               | 323           | 150<br>53.6%  | 101<br>32.7%  | 67<br>33.7%   | 53<br>20.9%   | 39<br>26.4%   | 35<br>10.3%   |
| wave3 |               |               | 357           | 159<br>55.5%  | 104<br>34.6%  | 67<br>35.6%   | 55<br>17.9%   | 47<br>14.5%   |
| wave4 |               |               |               | 364           | 185<br>49.2%  | 128<br>30.8%  | 96<br>25.0%   | 68<br>29.2%   |
| wave5 |               |               |               |               | 371           | 146<br>60.6%  | 101<br>30.8%  | 56<br>44.6%   |
| wave6 |               |               |               |               |               | 357           | 198<br>44.5%  | 127<br>35.9%  |
| wave7 |               |               |               |               |               |               | 432           | 148<br>65.7%  |
| wave8 |               |               |               |               |               |               |               | 484           |

上段：サンプル数  
下段：t-1時点に対するt時点の  
アトリション割合

斜線部：refreshments, 白部：panel sample

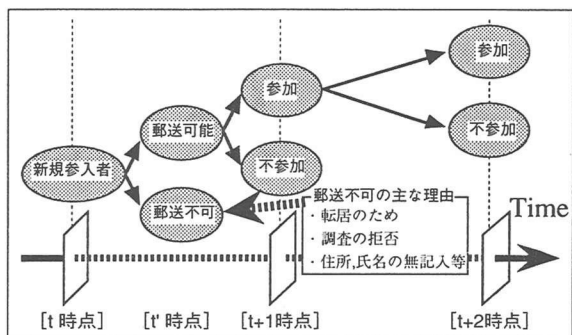


図-1 初期アトリションの概念

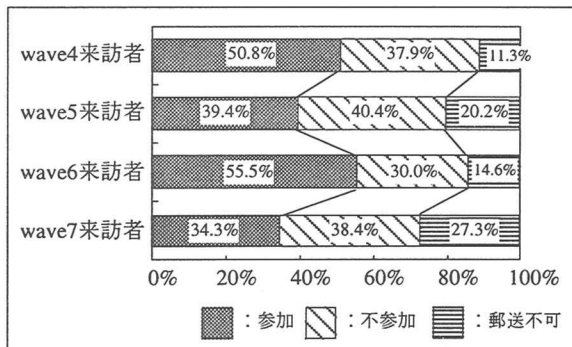


図-2 初期アトリションの割合

のパネル離脱は回避できないが、アンケート調査に個人情報 (住所、氏名) を提供することへの忌避感を少なくするためには、調査の簡易化や項目の縮小などの調査法の検討も必要であるといえよう。また、このような議論は、主にPuget Sound Panelやその他海外のパネル研究に詳しく紹介されており、例えば電話での事前承諾による調査実施、調査結果概要の紹介や調査依頼のレターを出し被験者との情報交換を緊密にすることでアトリションの割合が少なくなるなど、研究事例がいくつか報告さ

れている。

また、図-3は、wave6 来訪サンプルの年齢構成に着目し、次時点調査への参加・不参加・郵送不可別のサンプル構成を示したものである。来訪時点での構成比率に比べ、次時点すなわち wave7での参加者の構成比率が異なることがわかる。特に、20歳代の参加者数が、来訪時の半数以下になるのに対し、50歳以上は7割弱が次調査に参加している。これは、20歳代の郵送不可者が最も多いことから、若年層ほど調査から離脱する傾向が示されている。

ここでは、年齢構成の集計結果を取り上げたが、その他の要因に関してもこのような偏りが生じていることが推測でき、次章ではこのアトリションの構造を、その他の要因も含め明らかにしていく。

3. アトリションモデル（2項選択モデル）の構築

ここでは、参加・不参加を外的基準とする選択モデルを構築することにより、パネルアトリションの規定要因構造の記述をしていく。そこで、以下の3ケースを比較することでそれぞれの時点でのアトリションの違いを考慮していく。データセットは、調査票が比較的安定してきた wave4 に新規で調査に参入してきたサンプルを用いる。具体的には、図-4に示すように、

- (ケース1) 来訪 (wave4) のデータ
  - 1 時点後の参加・不参加
- (ケース2) 1 時点後 (wave5) のデータ
  - 2 時点後の参加・不参加
- (ケース3) 2 時点後 (wave6) のデータ
  - 3 時点後の参加・不参加

を定義した。ここで、ケース1は初期アトリション、そしてケース2、3はそれ以降のアトリションの規定要因構造をみるものである。

また、このアトリションモデルに用いる説明変数であるが、参加・不参加に関する基礎集計分析で、参加・不参加者間のサンプル構成に比較的差異の見られたものを取り上げている。具体的には、表-2に示した個人属性に関する6項目と対象SCに対しての利便性評価値およびダミー変数を用いる。そして、この推計結果を表-3に示す。ここで、表-

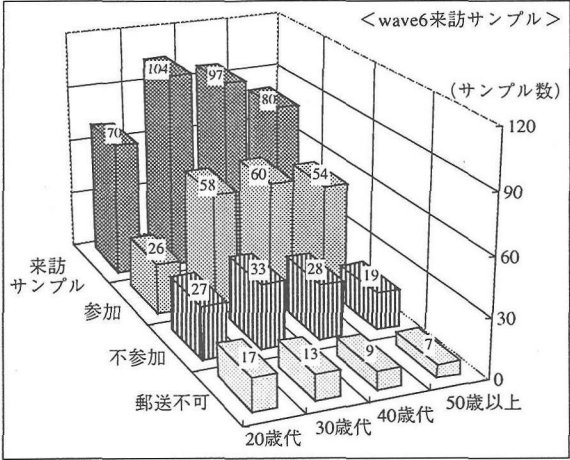


図-3 アトリション分類別年齢構成

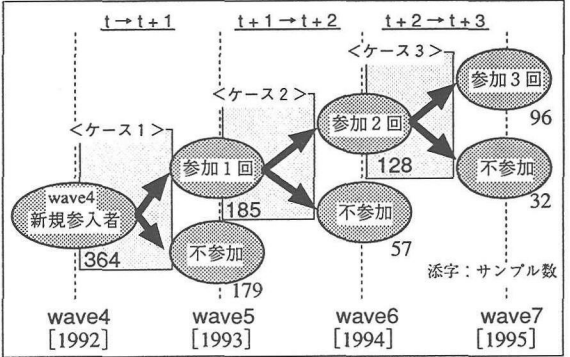


表-2 アトリションモデル説明変数

| 説明変数        | カテゴリ                                              | 説明変数                                          | カテゴリ         |
|-------------|---------------------------------------------------|-----------------------------------------------|--------------|
| 1. 性別 P     | 男性=0 女性=1                                         | 6. 利用交通手段 P                                   | 自家用車=1 その他=0 |
| 2. 年齢 P     | 30歳未満=1 30歳代=2<br>40歳代=3 50歳以上=4                  | 7. SC交通利便性 P                                  | 評価値          |
| 3. 子供の人数 P  | 18歳未満人数                                           | 8. SC活動利便性 P                                  | 評価値          |
| 4. 自家用車保有 P | 保有台数                                              | 9. 固有ダミー P                                    | 1            |
| 5. 居住年数 P   | 5年以下=1<br>6年～10年=2 11年～15年=3<br>16年～20年=4 21年以上=5 | 利便性の評価の定義<br>：調査票において10段階で評価する<br>P：参加側への固有変数 |              |

3と図-4においてパラメータ推計する際に用いるすべての説明変数の値が既知であるサンプル数のみを抽出したため、結果的にサンプル数が減少してしまった。

まず、全体的な適合度を示す尤度比 ( $\rho^2$  値) は3モデルとも小さく、特に時点が進んでいない段階の初期アトリションにおける適合度が悪い。これは、参加・不参加行動のランダム性にその原因があると考えられる。一方、時点が進めば、このような被験者は減少し、次第に適合度も高くなる傾向にある。

次にt値を見ていくと、各ケースで高い説明力を示す変数の組み合わせが異なってくる。まず、ケース1では「年齢」と「利用交通手段」が、ケース2では「年齢」、「子供の人数」、「居住年数」と「利便性評価」が、ケース3においては「性別」と「利便性評価」が高い説明力を持つ。

そして、個々のパラメータの符号に着目すると、年齢と子供の人数は正のパラメータであるので、年齢が高いほど、子供の人数が多いほど参加傾向にあり、これは時点を経ても変化しないことがわかる。また利便性評価の変数に関しては、t値は高いものの、その符号条件より評価が高いほど不参加傾向になるなど必ずしも十分な説明ができていない。

このように、モデルから得られた適合度は必ずしも良好とは言えないが、初期とそれ以降の時点におけるアトリションとは、今回の基礎集計結果からも、またこれまでの Attrition model の適合度及び推定されたパラメータ値の傾向からも、異なる構造をもつものと考えられる。しかし、時点の進行とともにこうしたアトリション構造の変化を追跡できているかについては、今回の分析からは必ずしも明確な結果を得たわけではない。

#### 4. 多時点間アトリションモデル：

オーダードロジットモデル

このモデルは、買物交通の発生回数や自動車の保有台数の選択など、序数的に並ぶ選択肢間の選択を連続した段階選択と仮定して、各段階の選択をロジットモデルを用いて説明するものである。ここでは、多時点にわたるパネル調査への参加・不参加の段階を、序数的に並ぶ選択肢を段階的に選択していくものとし、このオーダードロジットモデル (OL モデル) を適用していく。

ここで具体的な推計方法を述べる。まず、個人nのi番目の効用を $U_{i,n}$ とすると、i番目の序数を選択する確率 $P_n(i)$ は、0から順にiまで選択する確率 $\Pr(U_{1,n} > U_{0,n}) \cdots \Pr(U_{i,n} > U_{i-1,n})$ と、i+1番目以上が選択されない確率 $\Pr(U_{i+1,n} < U_{i,n})$ の積となる。

$$P_n(i) = \Pr(U_{i+1,n} < U_{i,n}) \prod_{k=1}^i \Pr(U_{k,n} > U_{k-1,n}) \quad (1)$$

表-3 アトリションモデル推計結果

| 説明変数     | ケース1     |       | ケース2    |       | ケース3    |       |
|----------|----------|-------|---------|-------|---------|-------|
|          | パラメータ    | t値    | パラメータ   | t値    | パラメータ   | t値    |
| 1.性別     | -0.102   | -0.34 | -0.224  | -0.51 | 0.840   | 1.54  |
| 2.年齢     | 0.244    | 1.70  | 0.286   | 1.21  | 0.311   | 0.94  |
| 3.子供の人数  | 0.031    | 0.25  | 0.495   | 2.21  | 0.246   | 0.87  |
| 4.自家用車保有 | -0.115   | -0.83 | 0.174   | 0.76  | 0.194   | 0.67  |
| 5.居住年数   | -0.025   | -0.27 | 0.280   | 1.91  | 0.090   | 0.47  |
| 6.利用交通手段 | -0.675   | -1.70 | -0.106  | -0.21 | 0.313   | 0.53  |
| 7.交通利便性  | 0.041    | 0.61  | -0.129  | -1.24 | 0.235   | 1.47  |
| 8.活動利便性  | -0.099   | -1.21 | -0.122  | -1.12 | -0.347  | -1.73 |
| 9.固有ダミー  | 0.915    | 1.08  | 0.556   | 0.47  | -0.680  | -0.38 |
| 初期尤度     | -187.028 |       | -95.920 |       | -63.465 |       |
| 最終尤度     | -181.452 |       | -87.603 |       | -58.915 |       |
| 尤度比      | 0.029    |       | 0.086   |       | 0.071   |       |
| 的中率      | 60.1%    |       | 71.6%   |       | 78.8%   |       |
| 有効サンプル数  | 271      |       | 155     |       | 118     |       |

95%有意水準  $t_{0.95}=1.96$  90%有意水準  $t_{0.90}=1.64$

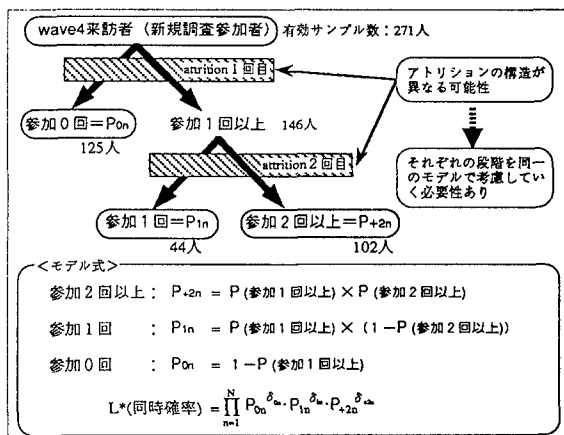


図-5 オーダードロジットモデルの概要

さらに、 $1 - P_{i+1,i,n} = \Pr(U_{i,n} > U_{i+1,n})$  とすれば、式(1)は次式のようにになる。

$$P_n(i) = (1 - P_{i+1,i,n}) \prod_{k=1}^i P_{k|k-1,n} \cdots (2)$$

ここで、各序数の選択確率 $P_{i+1,i,n}$ はi回か(i+1)回以上かの2肢選択を考えるBLモデルが適用できる。このときの、同時確率 $L^*$ を表現すると、

$$L^* = \prod_n (1 - P_{i+1,i,n}) \prod_{k=1}^i P_{k|k-1,n} \cdots (3)$$

となり、この $L^*$ を最大にするパラメータを推定していく。パラメータの推計方法は前述した最尤推定法を用いる。

このモデルを具体的に本研究に適用していくことになるが、本研究ではパネル調査に参加するか、

参加しないか、参加する場合1回だけか、2回以上か、という選択段階を設け、アトリションを記述していく。ここで、具体的なモデルの概要を図-5に示す。このモデルに導入する説明変数は、各選択段階において不参加の効用を0、参加の効用を個人属性、利便性評価と選択段階ごとのダミー変数を導入し定義している。また、それぞれの説明変数は表-2と同様のものを用いる。

まず第一に、簡便法の一つとして、各段階ごとのパラメータは変化させずに選択段階の違いを参加回数ごとのダミー変数で定義する方法を考え、そのパラメータ推計を行った。しかし、この方法によるパラメータ推計結果は、参加回数を表すダミー変数のみが有意となる以外は、説明力の高い変数が現れず、また尤度比も非常に小さい結果となった。そこで、各段階で異なるパラメータを設定することにした。今回においてモデルの構造は、参加回数が序数的であることに着目しているため、オーダードロジットを適用した。なお、Nested Logit Model は参加回数という選択肢に関していくつかのグループ分けがなされるが、参加回数に応じた束ね方によってログサム変数を定義する必要があり、実用性の観点から今回の適用は考えていない。

表-4はその結果を示す。まず、全体的な適合度を見ると、表-2のアトリションモデルにおける、時点が進んだケース(ケース2,3)とはほぼ同じで、必ずしも良好な結果ではない。しかし、符号条件については、各段階別に推計した結果と同様の傾向を示していること、また個々の説明変数の規定力(t値)も、各段階で有意なものが含まれていることから、本モデルでは、多時点にわたるアトリションの構造的特徴をある程度表現しているといえる。具体的に述べると、アトリション1回目では、「年齢」、「利用交通手段」、「参加1回目ダミー」の規定力が高く、アトリション2回目においては、「年齢」、「子供人数」、「居住年数」、「参加2回目ダミー」の規定力が高くなっており、それぞれの段階でアトリションの規定要因が異なることも示唆している。

表-4 オーダードロジットモデル推計結果

| 説明変数      |        | パラメータ    | t-値   |
|-----------|--------|----------|-------|
| アトリション1回目 | 性別     | -0.049   | -0.45 |
|           | 年齢     | 0.122    | 1.74  |
|           | 子供の人数  | 0.018    | 0.13  |
|           | 自家用車保有 | -0.057   | -0.81 |
|           | 居住年数   | -0.019   | -0.18 |
|           | 利用交通手段 | -0.335   | -4.84 |
|           | 交通利便性  | 0.025    | 0.42  |
|           | 活動利便性  | -0.048   | -0.84 |
| 参加1回目ダミー  |        | 0.450    | 6.46  |
| アトリション2回目 | 性別     | -0.110   | -1.58 |
|           | 年齢     | 0.164    | 2.46  |
|           | 子供の人数  | 0.259    | 3.78  |
|           | 自家用車保有 | 0.104    | 1.54  |
|           | 居住年数   | 0.172    | 2.65  |
|           | 利用交通手段 | -0.046   | -0.66 |
|           | 交通利便性  | -0.072   | -1.17 |
|           | 活動利便性  | -0.008   | -0.15 |
| 参加2回目ダミー  |        | 0.231    | 3.34  |
| 初期尤度      |        | -289.042 |       |
| 最終尤度      |        | -264.524 |       |
| 尤度比       |        | 0.085    |       |
| 的中率       |        | 73.4%    |       |
| 有効サンプル数   |        | 271      |       |

95%有意水準  $t_{0.95}=1.96$  90%有意水準  $t_{0.90}=1.64$

## 5. アトリションを考慮した買物場所選択モデル

パネルデータを扱う際には、前章まで述べてきりパネルアトリションによるサンプルの偏り(バイアス)の影響を考慮しなければならない。そこで、本研究では、バイアスを修正する方法としてWESML推定法(weighted exogenous sample maximum likelihood estimator)を適用した。

この推定法は、単純無作為標本抽出法(simple random sampling)より効率的なサンプリングを行える選択肢別標本抽出(choice based sampling)を適用するときの調査技法の一つで、この抽出法で得られたデータを前提とする最尤推定法としてManski,C.&S.Lerman (1977)<sup>8)</sup>によって提案された手法である。

WESML推定量は、母集団シェアとサンプル内シェアの比で表される重み  $w(g)$  を尤度関数式に掛け合わせたもので、次式で示される。

$$L^*(\alpha) = \sum_{g=1}^G \sum_{n=1}^{N_g} \sum_{i \in A_n} \delta_{in} w(g) \ln P(i|X_n, \alpha) \quad (4)$$

$$w(g) = Q(g)/H(g) \quad \dots (5)$$

Q (g) : グループ g の母集団中のシェア  
H (g) : グループ g のサンプル中のシェア

ここで、w (g) が 1 より大きいグループ、つまり母集団シェアよりも少なく抽出されているグループに対しては、一つのサンプルが推定量に与える影響を大きくし、逆に w (g) が 1 より小さいグループでは影響を小さくしようとするものである。ここで、式 (4) の最大化によって得られる推定量を WESML 推定量という。

この手法を、R.M.Pendyala & R.Kitamura (1992)<sup>9)</sup> は、アトリションバイアスを修正する手法として初めてパネル研究に適用した。また、佐々木・森川 (1996) はアトリションの離脱傾向をハザード関数を適用することによって定義し、そのバイアスの修正方法としてこの WESML 手法を適用し、目的地選択モデルとアトリションモデルを同時確率式で定式化<sup>9)</sup> している。

本研究においても、アトリションモデルとしては前章の離散型選択モデルを用いることにし、バイアス修正法としては、買物場所選択問題に対してこの WESML 手法を適用する。すなわち、この修正法における重みづけには、前章で構築したアトリションモデルのパラメータ推計結果を用いる。その基本的な考え方は、不参加傾向にあるサンプルについて重みをつけるというものである。具体的には、まず第 1 段階として、前章で得られたアトリションモデルのパラメータによって各個人の参加確率を求め、第 2 段階として参加確率の逆数を尤度関数に重みづけすることによって、消耗バイアスを修正していくものである。

すなわち、第一段階では、

$$W_n = 1/P_n \quad \dots \dots \dots (6)$$

W<sub>n</sub> : 個人 n のもつウエイト

$$P_n : \text{次調査への参加確率} = \frac{\exp(\theta Z_n)}{1 + \exp(\theta Z_n)} \quad (7)$$

θ : アトリションモデルのパラメータ

Z<sub>n</sub> : アトリションモデルの説明変数

次に、第 2 段階におけるアトリションバイアスを考慮した尤度関数は、式 (8) で定義される。また、多時点の場合の式 (7) は、前述の式 (2) で示される P<sub>n</sub> (i) (i: i 番目時点) を用いて各時点における参加確率を定義する。

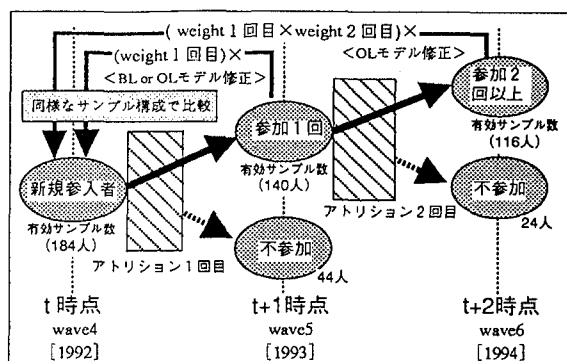


図-6 重みづけ買物場所選択モデル概要

$$L^* = \prod_{n=1}^N \left[ \frac{\exp(V_i)}{\sum_{j=1}^J \exp(V_j)} \right]^{W_n \delta_{in}} \quad \dots \dots \dots (8)$$

V<sub>i</sub>, V<sub>j</sub> : 効用関数、

δ<sub>in</sub> : 個人 n が選択肢 i を選択 = 1、  
そうでないとき 0

J : 選択肢数

以下では、この段階的なアトリションバイアス修正法を実際に適用し、買物場所選択モデルを構築していくが、図-6はその概要およびデータセットとの対応関係を示す。ただし、今回の推計方法が最良法であるかは、現段階では結論づけることはできなく、同時推定法などの他の修正法との比較が必要である。

図-6に示すように、まずここで1時点後の調査参加者（参加1回）について考える。このサンプルについては、1回のアトリションの影響を考慮し、その1回分の重みづけを行う。（この作業を行うことにより、仮想的に前段階（すなわち新規参入時点）と同様のサンプル構成となるようにする。）一方、参加2時点以上のサンプルについては、2回のアトリションを考慮、すなわち2回分の重みを乗じたものを用いる。この作業により、2時点後のサンプルについても新規参入時点と同様のサンプル構成で推計することができる。

次に、実際に考慮する重み (weight) であるが、参加1回目のサンプルには、前章で推計した2項選択モデル (BLモデル)、オーダードロジットモデル (OLモデル) 両ケースの推計パラメータを用いて算出する。一方、参加2回以上のサンプルについてはOLモデルで得られたパラメータを用い2回分のア

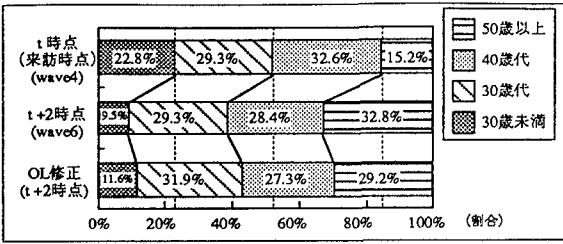


図-7 年齢構成の2時点後の修正状況

トリションの影響を考慮する。BLモデルは、アトリション2回目を推定する時点のサンプル構成(参加1回のサンプル)が、すでに1回目のアトリションの影響を受け、偏りが生じていると考えられるため、2時点後のバイアス修正には適用しない。

以上の修正法を用い、図-7には2時点後(t+2時点)の年齢構成の修正状況を示す。算出方法は、OLモデル(表-4)の推計結果から各個人の重みを求め、各カテゴリごとにt+2時点における修正サンプル数を算出した。ここで、OL修正とはOLモデルの推計パラメータから重みを算出し修正したものである。(ちなみに後述のBL修正とはBLモデルを修正に適用した場合を指す。)

この図から、アトリションバイアスを修正した2時点後の年齢構成において、30歳未満のサンプルの比率が高くなり、逆に50歳以上のサンプルの比率は低くなっており、年齢構成のバイアスはある程度は修正されているといえる。(ただし、中年層の30,40歳代が修正されていないが、これは年齢のカテゴリを線形として取り扱っていることに起因すると思われる。)

次に、この修正法を用いた1時点後(t+1時点)の買物場所選択モデルのパラメータ推計結果が表-5である。ここでは、調査参加から1時点後のサンプルに着目し、バイアス修正前後の推計結果を比較することで、アトリションバイアス修正法の有効性を明らかにしていく。ここでは、バイアス修正に用いる重み(weight)として、BLモデル修正(ケース1)とOLモデル修正の2種類を適用する。

まず、全体的な適合度は、重みづけ前後で変化しないが、重みづけすることにより見かけ上のサンプル数が増加するため、各変数の規定力が高くなっている。推計結果から、「自宅～買物場所距離」、「交通利便性」といった変数が高い規定力を持つ要因であることがわかる。また、規定力は小さい

表-5 買物場所選択モデル推計結果 (t+1時点)

| 説明変数      | t+1時点(Non-weight) |       | BLモデル修正 |       | OLモデル修正 |       |
|-----------|-------------------|-------|---------|-------|---------|-------|
|           | パラメータ             | t-値   | パラメータ   | t-値   | パラメータ   | t-値   |
| 子供人数      | 0.084             | 0.20  | -0.076  | -0.24 | -0.005  | -0.02 |
| 自家用車保有台数  | -0.014            | -0.11 | 0.028   | 0.31  | 0.009   | 0.10  |
| 活動目的      | 0.152             | 0.80  | 0.185   | 1.32  | 0.166   | 1.20  |
| 自宅～買物場所距離 | -0.272            | -3.30 | -0.282  | -4.75 | -0.275  | -4.64 |
| 交通利便性     | 0.317             | 3.08  | 0.284   | 3.92  | 0.302   | 4.10  |
| 活動利便性     | -0.034            | -0.35 | -0.032  | -0.45 | -0.031  | -0.45 |
| 初期尤度      | -93.801           |       | -93.801 |       | -93.801 |       |
| 最終尤度      | -77.691           |       | -77.854 |       | -77.730 |       |
| 尤度比       | 0.171             |       | 0.170   |       | 0.171   |       |
| 的中率       | 77.1%             |       | 75.7%   |       | 77.8%   |       |
| 有効サンプル数   | 140               |       | 140     |       | 140     |       |

95%有意水準  $t_{0.95}=1.96$  90%有意水準  $t_{0.90}=1.64$

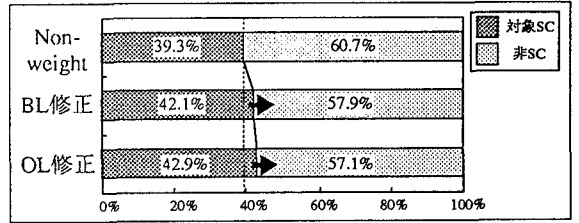


図-8 買物場所選択シェア (t+1時点)

表-6 買物場所選択モデル推計結果 (t+2時点)

| 説明変数      | t+2時点(Non-weight) |       | OLモデル修正 |       |
|-----------|-------------------|-------|---------|-------|
|           | パラメータ             | t-値   | パラメータ   | t-値   |
| 子供人数      | -0.589            | -1.16 | -0.671  | -2.16 |
| 自家用車保有台数  | -0.019            | -0.14 | -0.004  | -0.05 |
| 活動目的      | 0.574             | 2.55  | 0.541   | 3.91  |
| 自宅～買物場所距離 | -0.196            | -2.24 | -0.170  | -3.27 |
| 交通利便性     | 0.455             | 3.14  | 0.444   | 4.94  |
| 活動利便性     | 0.214             | 1.52  | 0.202   | 2.31  |
| 初期尤度      | -79.003           |       | -79.003 |       |
| 最終尤度      | -62.553           |       | -62.630 |       |
| 尤度比       | 0.208             |       | 0.207   |       |
| 的中率       | 70.7%             |       | 69.8%   |       |
| サンプル数     | 116               |       | 116     |       |

95%有意水準  $t_{0.95}=1.96$  90%有意水準  $t_{0.90}=1.64$

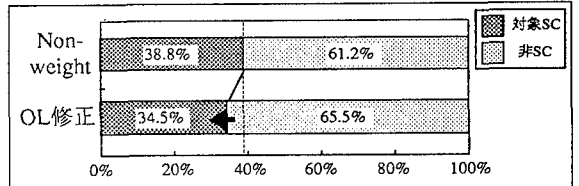


図-9 買物場所選択シェア (t+2時点)

が、「子供人数」と「自家用車保有台数」の符号が重みづけ前後で逆になっていることから、これらの変数に関してはバイアスがあったと推測でき、重みづけにより修正されたものと解釈できる。

また図-8は、パラメータ推計より算出された買物場所の選択シェアである。上段から、 $t+1$ 時点における修正前のモデル推計シェア、そして下段の2つが今回提案した修正法により推計した選択シェアである。これより、修正モデルにおいて、対象SCの選択シェアが推計値(修正前)より大きくなっている。BL修正およびOL修正による推定値は、1時点後の買物場所選択行動が来訪時点と同様のサンプル構成でなされた場合の選択シェアを示している。以上を踏まえ、対象SCを選択する層が1回目のアトリション、すなわち初期消耗の段階において、多く離脱している傾向があることを示唆している。

次に、2時点後( $t+2$ 時点)の買物場所選択モデルのパラメータ推計結果を表-6に示す。ここでは、バイアス修正に用いる重みとして、OLモデル修正を適用している。

$t+1$ 時点に比べ適合度は全体的に良好であり、「自宅～買物場所距離」、「交通利便性」は同様に高い規定力をもつ。そして、 $t+1$ 時点では比較的规定力の弱かった「子供人数」、「活動目的」および「活動利便性」も有意な規定力を示しているのが特徴的である。また、バイアスを修正することによってこの傾向が顕著に現れてくる。

このことから、来訪時点では距離と交通利便性だけで決められていた買物場所の選択構造は、時点が進むにつれ、子供の人数や活動目的といった個人属性や買物活動特性なども加わった形で、買物場所選択がなされる傾向をもつに至ることがわかった。

また図-9は、 $t+2$ 時点の買物場所の選択シェアである。これより、修正モデルにおいて対象SCの選択シェアが、修正前よりも下降していることから、2時点後の調査では、対象SCを選択している被験者がより多く残っていることを物語っている。

## 6. 結論

本研究では、多時点にわたるパネル調査のアトリションとそのバイアス修正法について検討した。まず、アトリションモデルでは初期とそれ以降のアトリションの構造の違いを明確にすることがで

きた。また、OLモデルを適用することにより、多時点にわたるパネルアトリションを記述できるようになった。

そして、このアトリションバイアスを修正した買物場所選択モデルでは、パネルサンプルがより母集団に近いサンプル構成をもつとの仮定のもとでパラメータ推計を行うことができ、より有効な時点間比較が可能となったといえる。

本研究では、基本的には参加・不参加と買物場所選択の意思決定は独立的に扱うことを前提としており、そのパラメータ推定は、この仮定のもとで段階的な推定を採用した。

ただし、参加・不参加と買物場所選択とが無関係であるとは言えず、こうした場合にはパラメータの同時推定も一つの方法であると言える。しかし、現実的には本分析でのアトリションモデルの推定結果によれば、この両者に強い相関があるとは考えられない。

以上、本研究では、パネル分析を行っていく際に問題となるアトリションについて実証的に検証してきた。今回はこのアトリションに伴うバイアスのみに着目してきたが、その他の諸バイアス(choice-based samplingのバイアス、無回答バイアス等)も当然存在しており、今後は消耗バイアス以外のバイアス問題にも目を向ける必要があると思われる。また、バイアスを修正することで得られる推計結果が個人の交通行動の意思決定構造を的確に表現しているのか、またそれをどのように検証していくかについても、今後の検討課題であろう。

### <参考文献>

- 1) K. Nishii & K. Kondo : Panel Analysis of Shopping Destination Choice Behavior in Japan , Paper presented at The First US Conference on Panels in Transportation, Arrowhead, CA., 1992.
- 2) 西井和夫, 古屋秀樹, 鈴木隆 : 甲府買物パネル調査データにおけるアトリション分析, 日本行動計量学会第22回大会発表論文抄録集, pp64-67, 1994.
- 3) 西井和夫, 近藤勝直, 古屋秀樹, 鈴木隆 : パネルアトリションを考慮した買物場所選択モデル, 土木計画学研究・論文集, No.12, pp389-396, 1995.
- 4) 杉恵頼寧, 藤原章正, 山根啓典 : 選好意識パネルデータに潜在する消耗バイアスの修正, 土木計画学研究・論文集, No.11, pp311-318, 1993.
- 5) 佐々木邦明, 森川高行, 杉山幸司 : パネルサンプルの初期消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, pp595-601, 1996.



- 6) 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, pp397-404, 1995.
- 7) 平成5・6年度科研総合研究 (A) (飯田恭敬)：交通計画におけるパネル調査データの適用性に関する研究, 1995.
- 8) C.Manski & S.Lerman：The Estimation of Choice Probability

from Choice Based Samples, *Econo-metrica*, Vol.45, pp.1 977-1 988, 1977.s

- 9) R. M. Pendyala & R. Kitamura：Weighting Methods for Choise Based Panels with Attrition, Paper presented at The First US Conference on Panels in Transportation, Arrowhead, CA., 1992.

---

#### 多時点パネルのアトリションバイアスとその修正法に関する研究：甲府買物パネルデータを用いて

西井和夫、近藤勝直、古屋秀樹、栃木秀典

本研究は、パネルデータ収集に関連する固有な問題であるパネルアトリションに関して、これまでに検討例の少なかった複数時点間に及ぶアトリションバイアスの修正法を提案し、その有効性を適用例を通じて検証した。具体的には、甲府買物パネルデータを用い、各ウェーブ間のアトリションを説明するモデルの同定化を行った。次いで序数型（非集計）ロジットモデルの考え方にもとづき多時点間のアトリション行動を表現できるモデルを構築するとともに、これを用いて買物場所選択行動におけるアトリションバイアスの修正法を検討した。その結果、初期アトリションとそれ以外との間の要因構造上の差異を示すことができたこと、またバイアス修正法の有効性も検討できた。

---

#### Panel Attrition and the Correcting Method for its Biases : Using Multi-Wave Kofu Shopper Panels

By Kazuo NISHII, Katsunao KONDO, Hideki FURUYA and Hidenori TOCHIGI

This paper aims to analyze the sample attrition problem peculiar to a panel survey, especially related to multi-wave attrition biases. The paper also intends to develop a correcting method for those attrition biases and empirically examines an applicability of this method to shopping destination choice behavior. First of all, attrition models are calibrated for several waves in Kofu shopper panels in order to identify the difference in decision structure between the first attrition and the following-step attritions. Secondly, the ordered logit model corresponding to multi-wave attrition process is formulated, and the weighting method for shopping destination choice behavior with those attrition biases is also developed. As a result, it is found from an empirical application that the proposed method is useful for correcting those biases.

---