

離散 - 連続モデルを用いた世帯の自動車複数保有・利用構造の分析*

A Discrete-Continuous Model for Analyzing Households' Multiple Vehicle Ownership and Their Car Usage*

小林 迪子**・福田 大輔***・兵藤 哲朗****・田中 倫英†

by Michiko KOBAYASHI**・Daisuke FUKUDA***・Tetsuro HYODO****・Tomohide TANAKA†

1. 研究の背景と目的

従来、自動車保有台数、運転免許保有者数、自動車走行台キロなどといった自動車に関連する統計値は、日本の経済成長とともに右肩上がりに伸びており、将来もそのような現象が続くものと考えられてきた。しかし、近年、少子高齢化、核家族化、ガソリン価格の高騰、地球温暖化といった社会経済情勢の様々な構造変化と共に、世帯構成や維持費等に影響される自動車市場の構造も大きく変化している。特に大きな変化として、2003年以降、自動車保有台数及び免許保有者数は増加する一方で、走行台キロが減少するという傾向に転じていること、また乗用車保有台数が横ばいであるのに対し、軽自動車保有台数が増加していることの2点が挙げられる。国全体の自動車交通量の予測は、道路の整備計画、有料道路の償還計画等に必要だが、社会情勢に大きな変化が生じている今、その方法にも大幅な修正が必要なことは言うまでもない。

従来の自動車交通量の予測は、将来推計人口、GDP、免許保有者数等をもとに行われてきた。これを新たに見直すにあたっては、近年の少子高齢化、世帯構造の変化、ガソリン価格の影響、技術革新による燃費向上などを考慮して世帯単位での分析を行い、走行台キロの減少の要因を解明することが必要となる。また、一般に軽自動車は普通乗用車に比べて走行距離が短いことから、普通乗用車から軽自動車への買い替えや軽自動車の買い増し等が行われた場合、全体として走行距離が減少すると予想される。これら2つの背景は互いに関連を持っており、世帯単位での自動車複数保有と利用の関連性を明らかにした上で自動車交通量の予測を行う必要があると考えられる¹⁾。

以上の問題意識のもと、本研究では、世帯・地域特性を考慮した自動車複数保有及び利用構造の分析を行うため、離散 - 連続モデルを用いた自動車保有 - 利用の同時決定モデルを構築し、平成11年・17年の道路交通

センサデータを適用して、その有効性を検証する。

2. 既往研究の整理

保有と利用を同時に扱う代表的手法は、効用理論に基づく離散 - 連続モデルである。代表例として、De Jong²⁾が挙げられるが、1台のみ保有する世帯と全く保有しない世帯を分析対象としており、複数保有世帯を対象とする研究は見られない。また、効用理論に基づく離散 - 連続モデルの適用には、世帯の所得に関する情報が必要であるが、道路交通センサには所得情報が含まれていないため、適用は難しいと考えられる。

これに対し、Fang³⁾は、複数保有の世帯も含めた上で、車種ごと(2種)の保有台数と走行距離について、それぞれの選択行動を効用関数として定式化し、Ordered Probitモデルを用いて保有台数を、またTobitモデルを用いて走行距離を決定する連立方程式体系のモデルを構築している。従来の離散 - 連続モデルの枠組とは大きく異なるが、複数保有と各保有車両の走行距離との関連性を極めて柔軟に表現することができる。本研究ではFangのBMOPT(Bayesian Multivariate Ordered Probit & Tobit)モデルを更に拡張した定式化を行う。

3. 離散 - 連続モデルの概要

(1) BSUROPTモデルの定式化

本研究では、FangのBMOPTモデルを拡張した、BSUROPT(Bayesian Seemingly Unrelated Regression, Ordered Probit & Tobit)モデルを構築する。Fangでは、全ての方程式で同一の説明変数が用いられる定式化となっている。これに対し、本研究では、方程式毎に異なる説明変数を設定することが可能なSUR(Seemingly Unrelated Regression)モデルの枠組みで新たな定式化を行う。ここでは更に、乗用車と軽自動車でも異なる説明変数を設定し、走行距離の効用関数に各車種の保有の有無を表すダミー変数を取り入れる。

乗用車、軽自動車それぞれの「保有台数 (y_{1i}, y_{2i}) 」及び各車種の保有台数を規定する「潜在効用 (y_{1i}^*, y_{2i}^*) 」、並びに、打ち切り(Censoring)を考慮しない各車種の「潜在的な平均走行距離 (y_{3i}^*, y_{4i}^*) 」及び「実平均走行距離 (y_{3i}, y_{4i}) 」について、連立方程式体系として定式化

* キーワーズ：離散 - 連続モデル，自動車保有・利用，世帯

** 学生会員 東京工業大学大学院理工学研究科土木工学専攻
(〒152-8552 東京都目黒区大岡山 2-12-1-M1-11)

TEL 03-5734-2577, FAX 03-5734-3578)

*** 正会員 東京工業大学大学院理工学研究科土木工学専攻

**** 正会員 東京海洋大学海洋工学部流通情報工学科

† 非会員 国土交通省道路局企画課道路経済調査室

を行う。ここで、走行距離とは、各世帯における、車種別の1台あたり1日平均走行距離を表す。

まず、各世帯の意思決定構造は、次の潜在変数モデルを用いた連立方程式体系として定式化される。

$$\begin{aligned} y_{1i}^* &= \mathbf{x}_{mi}'\beta_1 + \epsilon_{1i} \\ y_{2i}^* &= \mathbf{x}_{si}'\beta_2 + \epsilon_{2i} \\ y_{3i}^* &= \mathbf{x}_{mi}'\beta_3 + \epsilon_{3i} \\ y_{4i}^* &= \mathbf{x}_{si}'\beta_4 + \epsilon_{4i} \end{aligned}$$

ここで、 $i = 1, \dots, N$ は世帯、 $j = 1, \dots, 4$ は方程式番号、 $k = m, s$ はそれぞれ乗用車、軽自動車を表すラベル、 y_{ji}^* は潜在変数ベクトル、 \mathbf{x}_{ki} は世帯・地域・自動車特性等の説明変数ベクトル、 β_j は世帯・地域・自動車特性の未知パラメータベクトル、 ϵ_{ji} は誤差項を表す。ベクトル ϵ_j は次の多変量正規分布に従うものとする。

$$\epsilon_j \sim \text{i.i.d. } \mathcal{MVN}(\mathbf{0}, \Sigma)$$

ここで、 Σ は誤差項ベクトルの分散共分散行列である。

次に、自動車保有に関する測定方程式、つまり、乗用車・軽自動車の観測される保有台数とそれぞれの潜在効用との対応関係は以下のように表わされる。

$$y_{1i} = \begin{cases} 0 & \text{if } y_{1i}^* \leq \alpha_{11} \\ 1 & \text{if } \alpha_{11} < y_{1i}^* \leq \alpha_{12} \\ 2 & \text{if } \alpha_{12} < y_{1i}^* \leq \alpha_{13} \\ 3 & \text{if } \alpha_{13} < y_{1i}^* \end{cases}$$

$$y_{2i} = \begin{cases} 0 & \text{if } y_{2i}^* \leq \alpha_{21} \\ 1 & \text{if } \alpha_{21} < y_{2i}^* \leq \alpha_{22} \\ 2 & \text{if } \alpha_{22} < y_{2i}^* \leq \alpha_{23} \\ 3 & \text{if } \alpha_{23} < y_{2i}^* \end{cases}$$

ここで、 $\alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{21}, \alpha_{22}, \alpha_{23}$ は各潜在効用の閾値であり、Fang に従うと、標準累積密度関数 Φ を用いて、 $\alpha_{11} = \alpha_{21} = -\Phi^{-1}(1/4)$ 、 $\alpha_{12} = \alpha_{22} = 0$ 、 $\alpha_{13} = \alpha_{23} = \Phi^{-1}(1/4)$ という定数として与えることができる。このような Ordered Probit モデルの閾値設定方法の場合、最大・最小の閾値を定め、その他の閾値を推定するのが便利であることから⁴⁾、ここでは、 α_{12}, α_{22} をパラメータとして推定する。なお、道路交通センサスを用いた実証分析では、非保有世帯を分析の対象外としていることから、乗用車・軽自動車の台数が同時に0台になることはない($y_{1i} = 0$ かつ $y_{2i} = 0$ になることはない)という点に留意する必要がある。

最後に、走行距離に関する“測定方程式”については、各車種について保有台数が0台の場合、走行距離は必ず0になることから、Tobit モデルの理論を適用することで、以下の関係式を得ることができる。

$$y_{3i} = \begin{cases} y_{3i}^* & \text{if } y_{3i}^* \geq 0 \\ 0 & \text{if } y_{3i}^* < 0 \end{cases}$$

$$y_{4i} = \begin{cases} y_{4i}^* & \text{if } y_{4i}^* \geq 0 \\ 0 & \text{if } y_{4i}^* < 0 \end{cases}$$

(2) 推定方法

本研究では、ベイズ推定の1つであるマルコフ連鎖モンテカルロ法(以下、MCMC法)を用いて推定を行う。本研究における尤度関数は非常に複雑な構造であり、更に積分が閉じていないため、最尤法によるパラメータ推定は困難を極める。ベイズ推定を利用することで、多重積分を直接解かずに推定を行うことができ、計算コストや時間の短縮が可能になる。ここでは、MCMC法の中でも、条件付分布から標本を生成する Gibbs Sampling アルゴリズムを用いて推定を行う。

BSUROPT モデルにおいて確率変数として扱われる変数は、潜在効用 y_i^* 、未知パラメータベクトル β 、分散共分散行列 Σ 、閾値 α_{12}, α_{22} である。本研究では、以下のような Gibbs Sampling アルゴリズムに従い、これらの未知変数の標本を生成する。

Gibbs Sampling アルゴリズム

- (1) 初期値を設定する。
- (2) $\pi(\beta|\Sigma^{(t-1)}, \mathbf{y}_i^{*(t-1)}, \alpha)$ から $\beta|\Sigma, \mathbf{y}_i^*, \alpha$ を生成。
 $\pi(\Sigma|\beta^{(t)}, \mathbf{y}_i^{*(t-1)}, \alpha)$ から $\Sigma|\beta, \mathbf{y}_i^*, \alpha$ を生成。
 $\pi(\mathbf{y}_i^*|\beta^{(t)}, \Sigma^{(t)}, \mathbf{y}_i, \alpha)$ から $\mathbf{y}_i^*|\beta, \Sigma, \mathbf{y}_i, \alpha$ を生成。
 $\pi(\alpha|\beta^{(t)}, \Sigma^{(t)}, \mathbf{y}_i^{*(t)})$ から $\alpha|\beta^{(t)}, \Sigma^{(t)}, \mathbf{y}_i^{*(t)}$ を生成。
- (3) 収束するまで(2)を繰り返す。

各変数の事後分布はそれぞれ、 \mathbf{y}_i^* : 打ち切りのある一変量正規分布、 β : 多変量正規分布、 Σ : 逆ウィッシュャート分布、 α_{12}, α_{22} : 一様分布となる⁵⁾。

4. データの概要

本研究では、平成11年・17年の道路交通センサスデータを世帯単位で再集計したデータを用いる¹⁾。道路交通センサスを利用することで、住所、世帯人数といった世帯属性や、年齢、性別等の個人属性に加え、保有する自動車の台数、各車両の車種、1日の走行距離等の自動車特性を各世帯別に得ることができる。世帯数は2時点の合計で981,240件(平成11年:513,914件、平成17年:467,326件)であるが、乗用車と軽自動車を合わせて1~4台保有する世帯のみを分析対象とし(各車種については0~3台とする)、このうち無作為に抽出した20,000世帯(各時点で10,000世帯ずつ)をモデル推定に用いる。更に、道路交通センサスから得ることのできない地域別の社会経済特性(例えば、人口密度、ガソリン価格、平均所得、駅数など)については、民力⁶⁾などその他のデータベースより補完する。

説明変数と被説明変数(各車種の保有台数、1台当たり平均走行距離)の基礎集計結果を表-1, 2, 3に示す。但し、表-3の平均走行距離については、左側の値が全世帯の1台当たり平均走行距離の平均値、右側の値(カッコ内の値)が利用しない世帯(走行距離0)を

除いた1台当たり平均走行距離の平均値を表す。

表 - 4 推定結果

表 - 1 説明変数の基礎集計

説明変数	平均	標準偏差
世帯人数(人)	3.00	1.41
子供の人数(人/世帯)	0.53	0.89
免許保有者数(人/世帯)	2.03	0.96
人口密度(人/km ²)	2385	3323
道路実延長合計(km)	28913	16625
10月ガソリン価格(円)	114	17
労働者平均所得(円)	365059	42358
バス路線延長比率	0.60	0.19

バス路線延長比率：道路実延長に対するバス路線延長の比率

表 - 2 保有台数の基礎集計

		軽自動車台数				合計
		0	1	2	3	
乗用車台数	0	-	3151 (15.8)	229 (1.1)	25 (0.1)	3,405 (17.0)
	1	11,383 (56.9)	1,964 (9.8)	233 (1.2)	33 (0.2)	13,613 (68.1)
	2	1,918 (9.6)	488 (2.4)	91 (0.5)	-	2,497 (12.5)
	3	356 (1.8)	129 (0.6)	-	-	485 (2.4)
	合計	13,657 (68.3)	5,732 (28.6)	553 (2.8)	58 (0.3)	20,000 (100.0)

表 - 3 走行距離の基礎集計

説明変数	平均	標準偏差
平均保有台数(台/世帯)	1.35	0.64
乗用車平均保有台数(台/世帯)	1.00	0.63
軽自動車平均保有台数(台/世帯)	0.35	0.55
乗用車平均走行距離(km/台)[平日]	14.03	22.98
軽自動車平均走行距離(km/台)[平日]	4.94	13.45
乗用車平均走行距離(km/台)[休日]	14.54(29.37)	27.43(32.93)
軽自動車平均走行距離(km/台)[休日]	4.22(23.33)	14.21(25.88)

パラメータ	事後分布の平均	標準偏差
(1) 保有台数(乗用車)		
世帯人数(人)	0.0137**	0.00405
人口密度(対数)	-0.0478**	0.00045
道路実延長(km)	0.0427**	0.00310
ガソリン価格[10月](円)	-0.0218**	0.00319
労働者平均所得(円)	0.0461**	0.00367
免許保有者数(人)	0.0973**	0.00406
鉄道駅数(駅)	0.0220**	0.00353
バス路線延長比率	-0.0052	0.00290
閾値	0.0003**	6.26E-10
(2) 保有台数(軽自動車)		
世帯人数	-0.0200*	0.00963
人口密度(対数)	-0.1396**	0.00098
車庫証明の有無(0:無, 1:有)	0.0427**	0.00474
道路実延長	-0.0480**	0.00474
ガソリン価格[10月]	0.0467**	0.00489
労働者平均所得	0.0171**	0.00530
子供(18歳以下)の人数(人)	0.0505**	0.00711
免許保有者数	0.1307**	0.00772
閾値	0.0008**	2.71E-08
(3) 1台当たり平均走行距離(乗用車)		
世帯人数	1.1180**	0.33727
人口密度(対数)	-7.0061**	0.16845
道路実延長	0.6369*	0.25192
ガソリン価格[10月]	-1.2597**	0.26712
労働者平均所得	1.6946**	0.33238
免許保有者数	1.4722**	0.33316
鉄道駅数	0.4445	0.31522
バス路線延長比率	-0.0694	0.26002
保有の有無(0:無, 1:有)	56.1300**	1.21000
(4) 1台当たり平均走行距離(軽自動車)		
世帯人数	0.4481**	0.64013
人口密度(対数)	-9.0672	0.17111
車庫証明の有無	3.4896**	0.32941
道路実延長	-0.9815**	0.31481
ガソリン価格[10月]	-0.4622**	0.32271
労働者平均所得	3.5162	0.38862
子供(18歳以下)の人数	1.1709**	0.48076
免許保有者数	1.4775*	0.51666
保有の有無	68.5390**	1.24230

5. 推定結果

パラメータ推定プログラムはOxを用いて作成した。変数選択を行って最終的に得られた推定結果は、表-4の通りである。表中の* (***) は、95% (99%) 信頼区間に0が含まれないことを意味する。

まず、乗用車の保有台数については、人口密度(対数)、ガソリン価格(10月)のパラメータが負になっており、人口密度が高いほど、またガソリン価格が上昇するほど、乗用車を選択する可能性が下がることがわかる。反対に、世帯人数、労働者平均所得、免許保有者数等のパラメータは正になっており、世帯人数や世帯内の免許保有者数が多い場合や所得が高い場合には、乗用車を保有する傾向が高まることを示している。

一方、軽自動車の保有台数については、世帯人数、道路実延長のパラメータが負であり、乗用車と反対の結果になった。すなわち、世帯人数が多い場合や、道路実延長が長い地域では、軽自動車を選択されにくいことが示唆される。また、ガソリン価格のパラメータは正であり、ガソリン価格が上昇すると、世帯の車種選択が乗用車から軽自動車へと移行する可能性が示唆される。次に、走行距離について見ると、保有の有無のパラメータの値が非常に大きく、支配的であることがわかる。車種別では、乗用車の人口密度のパラメータが負

になっており、人口密度が高い地域、つまり都市部では乗用車の利用が少ないことがわかる。ガソリン価格のパラメータは両車種で負になっており、ガソリン価格が上昇すると車種に関係なく、利用は減少すると言える。以上、パラメータの推定結果についてはいずれも直観に合う妥当な結果となっている。

続いて表-5に、誤差項分散共分散行列の推定値を、表-6に、分散共分散行列の対角項を標準化(平均0、分散1)した相関行列を表す。表中の台数(乗/軽)はそれぞれ乗用車、軽自動車の保有台数を、また距離(乗/軽)はそれぞれ乗用車、軽自動車の平均走行距離を表す。分散共分散行列より、保有台数の分散・共分散項が0.1前後と非常に小さい一方で、走行距離の分散・共分散項については、1000を超えるものもあるなど、非常に大きい値を取っている。従って、走行距離についてはばらつきが大きく、モデルの説明力向上が今後の重要な課題となる。モデルの精度が低い原因としては、次の2点が考えられる。1つ目に、保有するが利用しない世帯と、保有しないため必然的に走行距離が0になる世帯を同等に扱っていることが挙げられる。道路交通センサスの性質上、調査実施日に自動車を利用しない世帯が多くあり、本来、これらを考慮した分析を行わなければならない。2つ目の原因としては、固定する閾

値の設定が考えられる。本研究で構築した BSUROPT モデルでは、0 台と 1 台の境界となる閾値と 2 台と 3 台の境界となる閾値を固定し、1 台と 2 台の境界となる閾値を推定している。予め与える閾値については、正規分布の 25% と 75% の場所に固定しており、これが実データの各保有台数の割合にフィットした閾値では必ずしもない可能性がある。このため、全ての閾値の推定を行ったが、閾値の範囲内に収まる乱数が生成されず、推定を行うことができなかった。これは、パラメータの識別問題が生じているためと考えられる。閾値の設定については、今後の課題として検討が必要である。

次に、相関行列を用いて、各車種の台数と走行距離との相関関係を見ると、乗用車と軽自動車の保有台数、乗用車の保有台数と軽自動車の走行距離、軽自動車の台数と乗用車の走行距離の各相関係数が負になった。異なる車種間の相関係数は負になると予想されることからおおそ妥当な結果である。乗用車と軽自動車の走行距離の相関係数は正になっているが、絶対値が非常に小さいことから、両者に有意な相関は存在していないと考えられる。

表 - 5 誤差項の分散共分散行列

	台数(乗)	台数(軽)	距離(乗)	距離(軽)
台数(乗)	0.1454	-0.0937	2.1643	-0.3436
台数(軽)	-0.0937	0.2625	-1.3011	1.0135
距離(乗)	2.1643	-1.3011	1026.6000	49.1800
距離(軽)	-0.3436	1.0135	49.1800	607.7300

表 - 6 誤差項の相関行列

	台数(乗)	台数(軽)	距離(乗)	距離(軽)
台数(乗)	1.0000	-0.4795	0.1771	-0.0365
台数(軽)	-0.4795	1.0000	-0.0793	0.0802
距離(乗)	0.1771	-0.0793	1.0000	0.0623
距離(軽)	-0.0365	0.0802	0.0623	1.0000

6. 感度分析

推定結果を用いて、ガソリン価格の変動による感度分析を行う。ここでは、ガソリン価格を現状より 10%、25%、50% 上昇させた場合に、世帯ごとの各車種の選択確率がどのように変化するかを分析した。表 - 7 は各世帯の選択確率の変化を平均した値であり、表 - 8 は 1 台当たり平均走行距離の変化を全世帯で平均した値を示している。分析結果より、ガソリン価格が上昇するにつれて、軽自動車を選択する確率が高くなり、反対に普通車の選択確率が下がることがわかる。従って、ガソリン価格の上昇により、普通車から軽自動車を保有する傾向に移り変わると言える。また、平均走行距離の変化より、ガソリン価格が上昇するほど乗用車の利用は減少し、反対に軽自動車の利用が増える傾向が読み取れる。

表 - 7 各車種の選択確率の変化

ガソリン価格の 変化	乗用車の選択確率の変化			
	$\Delta P(0 \text{ 台})$	$\Delta P(1 \text{ 台})$	$\Delta P(2 \text{ 台})$	$\Delta P(M=3)$
10%	0.0054	-7.0E-06	-0.0020	-0.0054
25%	0.0136	-0.0001	-0.0049	-0.0135
50%	0.0273	-0.0006	-0.0100	-0.0267
ガソリン価格の 変化	軽自動車の選択確率の変化			
	$\Delta P(0 \text{ 台})$	$\Delta P(1 \text{ 台})$	$\Delta P(2 \text{ 台})$	$\Delta P(3 \text{ 台})$
10%	-0.0117	0.0037	0.0044	0.0080
25%	-0.0293	0.0089	0.0111	0.0204
50%	-0.0590	0.0169	0.0225	0.0422

表 - 8 1 台当たり平均走行距離の変化

ガソリン価格の変化	乗用車走行距離の 平均変化 (km)	軽自動車走行距離の 平均変化 (km)
10%	-0.716	0.081
25%	-0.889	-0.005
50%	-3.06	0.671

7. まとめ

本研究では、世帯特性・地域特性を考慮した自動車複数保有及び利用構造の分析を行うことが可能な新たな離散一連続モデルを構築した。また、推定結果を用いて、ガソリン価格の変動による感度分析を行った結果、ガソリン価格の上昇に伴い、軽自動車を保有する世帯が増え、同時に軽自動車の利用も増えるということが分かった。モデルの推定結果はおおよそ妥当なものとなったが、走行距離に関する説明力が十分ではなく、その原因と考えられる閾値の設定等については今後更に検討が必要であると考えられる。また、都道府県別データによる補完などもモデルの精度に影響を与えていると考えられ、グループごとに共通の非観測要因を明示的に考慮するマルチレベルモデル等の適用による改善を今後行う必要がある。

参考文献

- 1) 小林迪子, 福田大輔, 兵藤哲朗, 田中倫英 (投稿中) 道路交通センサスデータを用いた世帯の自動車複数保有及び利用構造の分析, 土木計画学研究・論文集.
- 2) De Jong, GC (1990) An indirect utility model of car ownership and private car use, *European Economic Review*, Vol. 34, No. 5, pp. 971-985.
- 3) Fang, H.A. (2008) A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density, *Transportation Research Part B*, Vol. 42, pp. 736-758.
- 4) Nandram, B., Chen, M. (1996) Reparameterizing the generalized linear model to accelerate gibbs sampler convergence, *Journal of Statistical Computation and Simulation*, Vol. 54, pp. 129-144.
- 5) 中妻照雄 (2003) ファイナンスのための MCMC 法によるベイズ分析, 三菱経済研究所.
- 6) 朝日新聞出版 (編) (2008) 民力 CD-ROM 2008, 朝日新聞出版.