

個人選択行動モデルの最近の開発動向に関するレビュー*

A Brief Review on Recent Discrete Choice Models for Travel Behavior Analysis*

兵藤哲朗**・室町泰徳***

Tetsuro HYODO**・Yasunori MUROMACHI***

1. はじめに

1990年代に入り、非集計モデルは計算機能力向上に伴うシミュレーション法によるモデル推定の一般化⁴⁾、最尤推定を容易にするソフトの普及などを背景に、極めて多様なバリエーションを生み出してきた²⁾⁶⁾。本論文は、それら諸モデルを幾つかの分類軸で整理し、実用化の観点からも考察を試みた簡単なレビュー論文である。今後のモデル開発や実用場面への適用可能性に関する議論の場を提供できれば幸いである。

2. 離散選択モデルの開発動向

(1) 選択枝間の誤差相関

MNL(Multinomial Logit)モデルで表現できない「赤バス・青バス」問題に代表される選択枝間の類似性の相違を、選択枝間誤差相関で表現する方法は、NL(Nested Logit)モデルにより確立され、実用面でも広く普及している¹⁾⁷⁾。しかしNLでは同一選択枝のネスト間の重複が認められないため、同じ選択枝が単独で選択される場合と、乗り継ぎモードの一部をなす場合(例えば車とP&Rなど)、適切なツリー構造を特定化できない。ネスト間の選択枝の重複を可能とするモデルは、Williams¹²⁾で既に“Mixed Mode”を擁するモデルとして紹介されている。しかし1990年代に入り、選択枝相互の誤差相関を前提とするPCL(Paired Combinatorial Logit)モデル¹⁴⁾や、ネスト間の選択枝重複を考慮したNLモデルの拡張型としてのCNL(Cross-Nested Logit)モデル¹⁰⁾などが、実データへの推定結果を伴い紹介されるにつれ、手段選択以外にも多く適用されるようになった。これらのモデルは全てGEV(Generalized Extreme Value)モデル²¹⁾から導出されるが、先頃発表されたGNL(Generalized Nested Logit)モデル¹¹⁾も、GEVから得られるモデルであり、適切な条件を付与することにより、NLやPCL、CNLが導かれることから、それ

ら諸モデルを含む一般性を有している。GEVモデルの原式では、これらは下記の通り表される。

・PCL:

$$G(y_1, \dots, y_n) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n (y_i^{\mu_{ij}} + y_j^{\mu_{ij}})^{\mu_{ij}} \quad (1)$$

・CNL:

$$G(y_1, \dots, y_n) = \sum_m \left(\sum_{n \in C_m} \alpha_{nm} y_n \right)^{\mu_m} \quad (2)$$

・GNL:

$$G(y_1, \dots, y_n) = \sum_m \left(\sum_{n \in C_m} (\alpha_{nm} y_n)^{\mu_m} \right)^{\mu_m} \quad (3)$$

ここで α_{mn} は、 $\sum_m \alpha_{mn} = 1$ を満たす、選択枝が各ネストに帰属する重みを表す allocation パラメータであり、 μ_m はネスト m のスケールパラメータ(いわゆるログサムパラメータ)である。式より、PCLはGNLにおいて特定のネストをとりあげ、allocation パラメータを固定したモデル、CNLはスケールパラメータを同様に固定したモデルと解釈される。

また、選択枝間誤差相関が強く現れる適用場面としては、ネットワーク計算時の経路選択モデルが古くより認識され、様々なモデルが提案されている。経路選択の場合、重複経路長という明確な相関指標が算出可能なため、それを用いたモデルが多く開発された。MNP(Multinomial Probit)モデルの共分散を重複率で説明する構造化Probitモデル^{13),15),20)}、重複率をMNLモデルの共通変数として取り込むCL(C-Logit)モデル²⁾や、PSL(Path-Size Logit)モデル²⁾が挙げられよう。

(2) ネットワーク均衡における重複経路問題

選択枝間誤差相関の典型的な例の1つは、ネットワークにおける、経路選択問題の選択枝間誤差相関である^{16),18),21)}。これは(1)で述べた選択枝間の誤差相関のモデル群に含まれるが、ネットワーク均衡モデルとの関連性が深いため、あえて本論文では(1)と異質な性格を有するモデルとして整理する。経路選択問題においては、複数の経路選択枝が物理的に重複する場合が普通であり、これに伴う誤差相関を考慮することが課題となる。MNPモデルは重複経路選択問題に対処し得る手法の1つである。重複経路長という物理的で観測可能な相関指標により誤差構造を仮定した構造化Probitモデルに関しては、実証的な適用研究も進められてい

*キーワード: 交通行動分析、交通配分

**正員、工博、東京商船大学流通情報工学課程

(東京都江東区越中島2-1-6、

TEL03-5245-7386、hyodo@ipc.tosho-u.ac.jp)

***正員、博士(工学)、東京大学工学部付属総合試験所

(東京都文京区弥生2-11-16、

TEL03-5841-7732、ymuro@ut.tu-tokyo.ac.jp)

る²⁰⁾。

誤差構造に対する強い仮定を持つ MNL モデル、仮定は緩和されるが同一選択枝のネスト間の重複が認められない NL モデルは一般的経路選択問題への適用が困難である。これらに対し、式(1)~(3)のモデルは誤差構造に関する仮定がさらに緩和され、ネスト間の重複が認められるという点で重複経路選択問題への対処が可能となっている。今後は GNL に関するさらなる理論的展開と実証的な適用、MNP との比較などに関する研究が進められよう。

(3) 選択枝内の誤差構造化

MNL では IID、同一分散スケールを持つ Gumbel 分布が各選択枝の誤差項として仮定されるが、これを一般化・構造化するモデル群としては、選択別に分散スケールの相違を考慮する HEV(Hetero-scedastic Extreme Value)モデル²⁵⁾や、MNP においてパラメータ分散を前提とする Taste Variation モデル²⁴⁾、パラメータ分散を時間価値分布に特化した Logit モデルである時間価値分布モデル^{22,23)}やその一般型がある。パラメータの分散を考慮することは、誤差分散がその説明変数に比例することと等価であるため、後者のモデルも誤差分散の一般化とみなせる。MNL を基礎とした誤差構造の一般化は、MXL(Mixed Logit)モデル^{3,5,8)} (Hybrid Logit または Kernel Logit モデルとも呼ばれる) により記述可能であるため、近年、その整理が容易になりつつある。また、MNP の誤差分散共分散行列の適切な設定により、これらのモデルはいずれも表現可能である。

(4) 繰り返しデータの誤差相関

選択枝毎の誤差項がサンプル内で相関を持つ、または選択枝間で誤差分散が異なる場合を総称して、非観測異質性と呼ぶ。後者については前節で述べた対策が可能だが、前者については複数サンプル間の尤度関数を構成するする必要が生じる。具体的には、SP データや、SP データと RP データの統合利用など、同一個人の複数サンプルを用いるケースで生じる特徴だが、これまで MPL(Mass-Point Logit)モデル^{26,27)}や、MD(Mixing Distribution)モデル²⁷⁾ (MXL と同一のモデル)、SP・RP データ間の系列相関を考慮したモデル²⁸⁾などが開発・適用されている。MXL は分析条件を絞ることができるパラメトリック手法であるが、説明力からは自由度の高いノンパラメトリック手法である MPL が優ることが報告されている。なお、図-1 は本研究で言及したモデル間の位置付けを示している。

3. 実用性からみた離散選択モデルの特徴

(1) 選択枝間の誤差相関

NL はこれまでも数多くの適用がなされ、選択枝間誤差相関を扱うモデルとして極めて一般的となった。更なる一般化を目指した PCL や CNL、GNL は、その一般化故に実用化の局面に限られる可能性がある。選択枝の組合せを限定した PCL はまだしも、例えば3選択枝で2層までのツリー構造を考えたとき、GNL では最大 109 通りのモデル構造の組合せが存在する(無論、その中に MNL、NL、PCL、CNL が含まれる)。表-1 はわが国の都市間交通データ(航

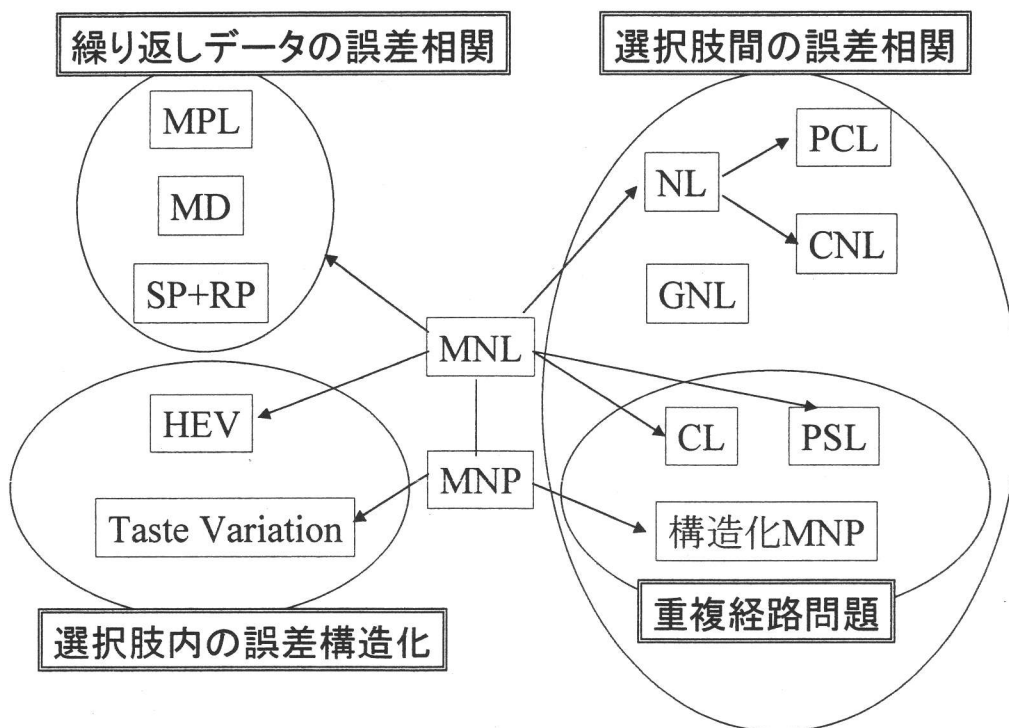


図-1 本研究で言及したモデル間の位置づけ
(なお MXL はいずれのモデル群も近似可能)

表-1 選択肢間の誤差相関を考慮したモデル [航空(A)、鉄道(R)、車(C)] (()内は t 値)

	MNL	NL	PCL	GNL
時間(G)	-1.144 (28.2)	-0.837 (19.7)	-0.829 (17.0)	-0.660 (13.0)
費用(G)	-3.060 (3.3)	-3.490 (5.7)	-1.698 (5.0)	-1.753 (19.3)
定数項(A)	-2.207 (18.8)	-1.083 (7.6)	-1.258 (13.8)	-0.989 (6.8)
定数項(R)	-0.990 (12.9)	-0.326 (3.4)	-0.876 (10.3)	-0.174 (1.9)
μ_{AR}		0.526 (12.3)	0.677	0.153
μ_{RC}			0.336	0.616
μ_{AC}			0.809	
$\alpha_{A(A)}$				0.164
$\alpha_{A(AR)}$				0.836
$\alpha_{R(AR)}$				0.527
$\alpha_{R(RC)}$				0.473
初期尤度	-3002.5	-3002.5	-3002.5	-3002.5
最終尤度	-2229.6	-2195.2	-2183.8	-2153.8
$\bar{\rho}^2$	0.2569	0.2683	0.2718	0.2815
サンプル数	3000	3000	3000	3000

時間は 100 分、費用は 10 万円単位。t 値記載のないものは制約条件を満たすためにロジット変換してパラメータ推定をしたことによる。「G」は共通変数を表す。また、 $\sum_i \alpha_i = 1, \alpha_i > 0$ である。

空・鉄道・車の3肢)への適用例であり、図-2はその際に仮定した GNL のツリー構造であるが、モデル構造が複雑になるほど説明力は高くなるものの、GNL におけるツリー構造を特定化するのは先験的情報がない限り困難である。むしろ mixed mode のような、選択肢間の相互包含関係が明確な対象への適用が望まれよう。

また、重複経路を明示的に扱うモデルである構造化 MNP、CL、PSL そして MXL については、その比較特性分析は構造化 MNP と MXL との比較検討に止まっており、より広範な分析が必要な段階にある。実用性から見れば、最も簡便な式型を有し、推定も容易な CL が優位であると思われるが、代替案の設定方法によっては CL の「共通項」(commonality factor)に係るパラメータが符号条件を満たさない場合もあり、必ずしも CL が万能とは云えない。決して新しいテーマではないが、代替案設定方法と、モデル推定結果との関連性分析は、次節の問題にも大きく関わる、未だ十分な解決案に至らない課題であろう。

(2) ネットワーク均衡における重複経路問題

経路選択問題に関する実用上の問題は、大きく経路選択モデルの推定と利用者均衡配分の側面に分けることができる。前者に関しては、特定の OD に関して代替的な経路選択肢集合が特定できれば、物理的で観測可能な重複経路長を、PCL の場合はスケールパラメータに、CNL の場合は allocation パラメータに反映することが可能である。しかし、依然として妥当な経路選択肢集合の特定手法が予め必要であり、また、モデルのパラメータ推定における求解の安定性も問題となる。

後者に関して、Prashker et al¹⁷⁾は PCL と CNL を前提とした利用者均衡配分と同値の最適化問題を定式化している。

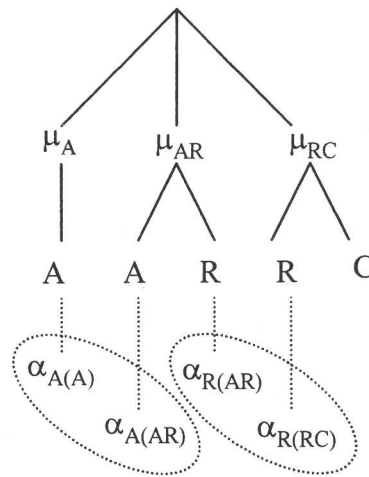


図-2 表-1で仮定した GNL のツリー構造

GNL に対する最適化問題は、CNL に対する問題のスケールパラメータ制約の緩和というごくわずかな変更により書き表せる(式(4))。逆に PCL に対する最適化問題は、ネストをペアワイズにした上で allocation パラメータを固定した問題となる。

利用者均衡配分を行うためには、さらに OD 需要をネットワークにローディングし、均衡計算を行う必要がある。CNL に対する均衡計算の具体的な方法は Vovsha et al¹⁹⁾が示している。なお、OD 需要をネットワークにローディングする際に妥当な経路選択肢集合の特定が必要となる点は、経路選択モデルの推定の際と全く同様である。また、簡易ネットワークを用いた Vovsha, P. et al による計算例では、混雑が激しくなるほど重複経路選択による MNL 配分と CNL 配分との差が縮小しており、激しい混雑が生じているネットワークでは MNL の重複経路問題が結果的に緩和される可

能性が示唆されている。

$$\min Z = Z_1 + Z_2 + Z_3 \quad (4)$$

$$Z_1 = \sum_a \int_0^{x_a} c_a(\omega) d\omega$$

$$Z_2 = \frac{\mu_m}{\theta} \sum_{rs} \sum_m \sum_k f_{mk}^{rs} \ln \frac{f_{mk}^{rs}}{(\alpha_{mk}^{rs})^{1/\mu_m}}$$

$$Z_3 = \frac{1 - \mu_m}{\theta} \sum_{rs} \sum_m \left(\sum_k f_{mk}^{rs} \right) \ln \left(\sum_k f_{mk}^{rs} \right)$$

$$s.t. \sum_m \sum_k f_{mk}^{rs} = q^{rs} \quad \forall r, s$$

$$f_{mk}^{rs} \geq 0 \quad \forall m, k, r, s$$

where

$$f_{mk}^{rs} = \text{OD}(r \rightarrow s), \text{ ネスト } m, \text{ 経路 } k \text{ の交通量}$$

$$q^{rs} = \text{OD}(r \rightarrow s) \text{ の交通需要}$$

$$c_a = \text{リンク } a \text{ のコスト}$$

$$x_a = \text{リンク } a \text{ の交通量}$$

$$\alpha_{mk}^{rs} = \text{OD}(r \rightarrow s), \text{ ネスト } m, \text{ 経路 } k \text{ の allocation パラメータ}$$

$$\theta = \text{dispersion 係数}$$

$$\mu_m = \text{ネスト } m \text{ のスケールパラメータ}$$

(3) 選択枝内の誤差構造化

交通手段選択モデル中の所要時間パラメータを分散を伴う random coefficient として推定すると、その分散が時間価値の分布と見なせる。誤差の構造化の一例であるが、その適用可能性は高いと思われる。図-3は、表-1と同じ都市間データを用いて得られた MXL による時間価値分布モデルによる感度分析例である。得られたパラメータを用いて、平均的な LOS を仮定し、航空の所要時間を 20%減少させると同時に航空費用を 20%増加させたときの時間価値別の logsum 値の増分を示している。時間価値分布を仮定しない場合は、logsum 値は全時間価値において同一の値をとるが、時間価値分布が存在する場合、時間価値が高い個人の logsum 増分が大きく、逆に時間価値が小さい個人ではその増分がないことが理解されよう。これは例えば、ロードプライシングなど、高サービス・高負担の政策に対しては、時間価値の相違に基づき賛否が大きく分かれることを端的に表しているといえよう。時間価値の相違に基づく公平性の評価をシステムティックに行うツールとしての適用性が期待される。

(4) 繰り返しデータの誤差相関

個人の繰り返しデータを用いた手段選択モデルにおいて、非観測異質性を考慮すると説明力が飛躍的に向上することが多い。SP 調査の重要性が増しつつある現在、同手法が標準的ツールとして実用上用いられることが望まれるが、盲

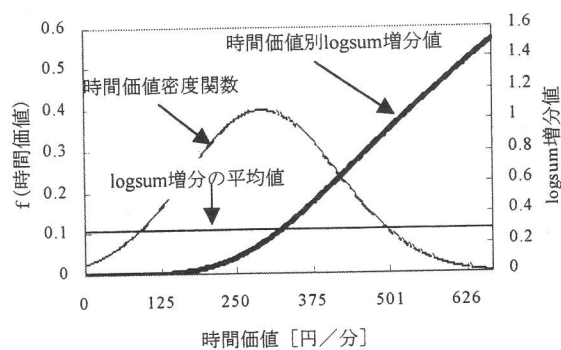


図-3 時間価値別の LOS 変化に対する logsum 値の相違

目的に適用するのではなく、非観測部分を減じるための説明変数の追加・工夫も必要である。また、非観測異質性を排除する究極の方法は、サンプル毎に定数項を推定することであり、MPL や MD、MXL などは、その定数項の確率的分布形状を推定することに他ならない。このことから容易に想像されるように、同質なサンプルをまとめる、いわゆるセグメント化も異質性を減じる重要な手法である。実用的には、適切なセグメント指標による古典的セグメント別モデル構築も、非観測異質性を克服する強力な一方法であることに留意すべきであろう。

4. おわりに

本論文では、主に 1990 年代に開発・提案された交通行動モデルの整理を目的に、選択枝間の誤差相関、重複経路問題、選択枝内の誤差構造化、非観測異質性の分類軸に沿った紹介を試みた。1970 年代より理論面、実証面から研究が進められてきた MNL、MNP の理解が進むにつれて、その限界が少しずつ明らかになり、MXL や GNL といったより柔軟性が高く、包括的なモデルの開発に至っているというのが現状であると思われる。

本論文でとりあげた分類軸に関する課題も多々残っていると考えられるが、その一方で柔軟性が高く、包括的なモデル開発に向かうことにより、必然的に生じてくる課題にも留意しなければならない。まず、本文中でも触れたように、柔軟性があるモデルがゆえに最良のモデル構造を特定することが難しく、モデル構造に関する分析者の判断の余地が増えることとなる。少なくとも初期段階で適合度の高いモデル構造が特定化できればよいが、1990 年前後に NL 以外の新たな GEV モデルが提案された時にも、実証的適合度が悪く、展開が進まなかった例もある。パラメータ推定における解の安定性の課題と共に今後の大きな課題となる。最良のモデルを判別する統計的手法の開発が必要である。

また、モデルの柔軟性の向上はパラメータ数の増加をもたらすことが多い。TDM などを念頭において、既存の交通施設の利用状況を詳細に検討する上では、柔軟性の向上が個人の交通選択行動の検討に幅を与え、大きな利益をもたらすことが予想される。反面、新規の交通施設の利用状況

を予測する上では、柔軟性の向上以外の目的をもたない解釈の困難なパラメータを新規に設定することが困難となろう。モデルに柔軟性を要求する個人の交通選択行動上の本質的な意味に関して、再度検討する必要があるのではないかと思われる。

最後に、本論文は個人選択行動モデルをランダム効用モデルを中心に開発状況をレビューしたが、必ずしもランダム効用モデル全体を網羅したわけではないし、ランダム効用理論に基づくモデルが個人選択行動モデルの全てであるというわけでもない。むしろ、これらのモデルは個人の交通選択行動を表現するモデル群のほんの一部であるという認識の方がもっともらしい。例えば、GNL は経路選択モデルにおける重複経路問題の解決に有効かも知れないが、紙面上でよく検討される典型的なネットワークはともかく、現実の都市内道路ネットワークを利用者が重複経路として認識しているかどうかは検討の余地がある。経路の認識自体、経路選択行動と不可分の関係にあると考えられ、このような点に関しては本論文でとりあげたモデル群のみで対処するのは困難であろう。

個人選択行動モデルの開発動向レビューを今行う意義の一つは、GEV 系で最も一般化された GNL が提案されたことである。また、ここでは触れなかったが、他には、ランダム効用モデルとして柔軟な構造を持つ MXL モデルの適用範囲の広さもレビューを行う契機となった⁵⁾⁸⁾。後者については、1970 年代以来様々な努力が傾注されてきたランダム効用モデル⁹⁾の根幹に関わるテーマであり、今後さらなる議論が進められればと思っている。

<参考文献>

[非集計行動モデル全般]

- 1) Ben-Akiva, M. and Lerman, S.R. (1985): *Discrete Choice Analysis*, MIT Press
- 2) Ben-Akiva, M. and Bierlaire, M. (1999): Discrete choice methods and their applications to short term travel decisions, in *Handbook of Transportation Science*, chap.2, Kluwer Academic Publishers
- 3) Brownstone, D. and Train, K. (1999): Forecasting new product penetration with flexible substitution patterns, *J. of Econometrics*, Vol.89, pp.109-129
- 4) McFadden, D. (1989): A method of simulated moments for estimation of discrete response models without numerical integration, *Econometrica*, Vol.57, No.5, pp.995-1026
- 5) McFadden, D. and Train, K. (2000): Mixed MNL models of discrete choice, *Journal of Applied Econometrics* (forthcoming)
- 6) McFadden, D. (2000): Disaggregate behavioral travel demand's RUM side -- A 30-year retrospective --, the 9th IATBR, Gold Coast, Australia
- 7) 原田 (1985): Nested Logit モデルの理論と適用に関する研究のレビュー、土木学会論文集 No.353, pp.33-42
- 8) 兵藤、章 (2000): Mixed Logit モデルの汎用性に着目した特性比較分析、土木学会論文集, No.660/ IV-49, pp.89-99
- 9) 屋井 (1986): 非集計行動モデルによる交通需要予測手法、交通と統計, No.16

[(1) 選択肢間の誤差相関に関するもの]

- 10) Vovsha, P. (1997): Application of Cross-Nested Logit model to mode choice in the Tel-Aviv, Metropolitan Area, TRR 1607, pp.6-15
- 11) Wen, C. and Koppelman, F.S. (2000): The Generalized Nested Logit Model, presented at the 79th TRR, paper number 00054312)
- 12) Williams, H.C.W.L. (1977): On the formation of travel demand models and economic evaluation measures of user benefit, *Environment and Planning A*, 1977, pp.285-344
- 13) 清水、屋井、坂井 (1998): 鉄道経路選択モデルにおける選択肢間の類似性の表現方法、土木計画学研究・講演集, No.21(1), pp.459-460
- 14) 藤原、杉恵、岡村、神田 (1999): 選択肢間の類似性を考慮した時刻選択モデル、土木計画学研究・講演集, No.22(1), pp.401-404
- 15) 屋井、中川、石塚 (1998): シミュレーション法による構造化プロビットモデルの推定特性、土木学会論文集, No.604/ IV-41, pp.11-21

[(2) ネットワーク均衡における重複経路問題に関するもの]

- 16) Bell, M.G.H. and Iida, Y. (1997) *Transportation network analysis*, Wiley
- 17) Prashker, J.N. and Bekhor, S. (1999) Stochastic user-equilibrium formulation for extended-logit assignment models, TRR1676, pp.145-152
- 18) Sheffi, Y. (1985): *Urban transportation networks*, Prentice-Hall
- 19) Vovsha, P. and Bekhor, S. (1998) Link-nested logit model of route choice : overcoming route overlapping problem, TRR1645, pp.133-142
- 20) Yai, T., Iwakura, S. and Morichi, S. (1997) Multinomial probit with structured covariance for route choice behavior, *Transportation Research*, Vol.31B, pp.133-152
- 21) 土木学会 (1998): 交通ネットワークの均衡分析、土木学会

[(3) 選択肢内の誤差構造化に関するもの]

- 22) Alam, M.J., Ohta, K. and Harata, N. (1996): Random distribution of value of time in discrete choice model for economic analysis of transportation projects, 土木学会 第 51 回年講, pp.378-379
- 23) Ben-Akiva, M., Bulduc, D. and Bradley, M. (1993): Estimation of travel choice models with randomly distribution values of time, *Transportation Research Record*, 1413, pp.88-97
- 24) Train, K. (1986): *Qualitative Choice Analysis*, MIT press
- 25) 福田、森地 (1999): 観光目的地選択行動に対する精緻化された個人選択モデルの適用可能性の検討、土木計画学研究・講演集, No.22(2), pp.655-658

[(4) 繰り返しデータの誤差構造化に関するもの]

- 26) 張 (1995): 異質性を考慮した交通需要モデルのダイナミック分析、日交研シリーズ A-186
- 27) 西井、北村、近藤、弦間 (1995): 観測されていない異質性を考慮した繰り返しデータによるパラメータ推定法: Mass Point Model と Mixing Distribution Model, 土木学会論文集, No.506, pp.25-33
- 28) 森川、山田 (1993): 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, pp.11-18

個人選択行動モデルの最近の開発動向に関するレビュー*

兵藤哲朗**・室町泰徳***

本論文では、主に 1990 年代に開発・提案された交通行動モデルの整理を目的に、選択肢間の誤差相関、重複経路問題、選択肢内の誤差構造化、非観測異質性（繰り返しデータの誤差相関）の分類軸に沿った紹介を試みた。1970 年代より理論面、実証面から研究が進められてきた MNL、MNP の理解が進むにつれて、その限界が少しずつ明らかになり、MXL や GNL といったより柔軟性が高く、包括的なモデルの開発に至っているというのが現状であると思われる。柔軟性の高さは個人の選択行動理解を広げる一方、最良なモデル構造の特定を困難にする面があり、モデル構造の解釈と改良に資する統計的手法の開発が必要である。

A Brief Review on Recent Discrete Choice Models for Travel Behavior Analysis *

By Tetsuro HYODO**・Yasunori MUROMACHI***

In this paper, we briefly review on recent developments in discrete choice models for travel behavior analysis. We summarize theoretical issues and practical concerns of new models in terms of “independence of irrelevant alternatives”(IIA) property, overlapping route choice, error structure among alternatives and unobserved heterogeneity. Extensive studies on Multinomial Logit/Probit Models (MNL/MNP) since 1970s have revealed their limitations, and lead to more flexible and comprehensive models such as Mixed Logit Models (MXL) and Generalized Nested Models (GNL). We can extend our scope of behavior analysis by those models, but the flexibility of model structure will leave much room of analyst's judgment.
