

東京大学都市工学科 学生員 原田 昇  
正員 太田 勝敏

1. 各種非集計モデルの検討と比較

非集計モデルとは、データをゾーンレベルの集計としてではなく、行動単位である個人レベルのままで分析するモデルである。特に、個人の意志決定プロセスを効用最大化理論を前提として説明する非集計モデルを、非集計行動モデルと呼んで区別する。効用関数のうち確率的に変動する部分の分布形の相違により、種々の行動モデルが導かれる。独立のWeibull分布、正規分布、矩形分布に対して、選択確率について、ロジットモデル、プロビットモデル、線型モデルが得られる。<sup>1)</sup>

ここでは、非集計行動モデルのロジットモデルと線型モデル、非集計モデルの数量化Ⅱ類と判別分析の4モデルの比較検討を行なう。分析の対象は、デパートへの買物交通の手段選択であり、サンプル数は113である。<sup>2)</sup>

【現状再現性のチェック】 4モデル別の有効な変数と、代表的な変数組によるキャリブレーションデータの現状再現性の比較である。変数の一覧を表1に示す。有効な変数の順位は、4モデルともに大差なく、SPN, AVAILA, HOYUDAI, RXXII, INCOME, DD2であった。

代表例として、定数項を含めて5変数のモデル、A [SPN, AVAILA, HOYUDAI, RXXII] と B [SPN, HOYUDAI, RXXII, DD2]

表1: 変数の一覧表

変数名	グループ名	サンプル-1 (N=57)		サンプル-2 (N=56)		サンプル-1+2 (N=113)	
		平均	標準偏差	平均	標準偏差	平均	標準偏差
DM1	利用手段の車	0.561	0.501	0.554	0.502	0.558	0.499
WN10	相対的利用頻度	0.556	0.316	0.557	0.316	0.557	0.315
LWN	$\ln(\frac{WN10}{1-WN10})$	0.283	1.690	0.294	1.706	0.289	1.688
HOYUDAI	保有台数/世帯	1.018	0.231	1.125	0.429	1.071	0.346
AVAILA	車の利用可能性	0.667	0.348	0.646	0.350	0.657	0.347
SPN	買物人数	2.228	1.240	2.196	1.197	2.212	1.213
INCOME	年収入(万円)	3.895	1.160	3.893	1.201	3.894	1.175
DD2	買物日	0.456	0.503	0.446	0.502	0.451	0.500
XX10	所要金額差	-71.987	132.530	-44.375	110.307	-58.303	122.269
XX11	所要時間差	1.049	10.498	2.750	11.471	1.892	10.975
GCOST	一般化費用	-47.803	269.080	19.013	307.654	-14.692	289.497
RXXII	XXIIの変形	-0.023	0.413	0.050	0.455	0.013	0.434
DXXII		0.053	0.789	0.143	0.749	0.097	0.767

\* INCOME = 1. 120万円未満 4. 480万円未満 \*GCOST = XX10 + 23.05 \* XX11 + DXXII = -1 XX11<sup>65</sup>  
2. 240 5. 600万円未満 \*RXXII = XX11 0 -55XXII<sup>65</sup>  
3. 360 6. 600万円以上 \*RXXII = 至利便手段、所要時間 1 55XXII

表2: 現状再現性の比較

モデルタイプ	適中率(%)	統計的指標	使用した指標
ロジット-A	89.38%	$\bar{P}^2 = 0.5596$ $R_c^2 = 0.6182$	$\tau$ 値, $k^2$ 値, $\bar{P}^2$ 値, $R_c^2$ 値
ロジット-B	89.38%	$\bar{P}^2 = 0.5335$ $R_c^2 = 0.6069$	
線型-A	86.73%	$R_c^2 = 0.5741$	F比, $R_c^2$ 値
線型-B	88.50%	$R_c^2 = 0.5603$	
教Ⅱ類-A	87.61%	$k^2 = 0.7562$	$k^2$ 値(相関比)
教Ⅱ類-B	88.50%	$k^2 = 0.7409$	
判別-A	84.96%	$\Lambda = 0.411$	Wilks's $\Lambda$ 等
判別-B	86.73%	$\Lambda = 0.424$	

の2ケースの分析結果を表2に示す。結果を比較する適切な指標はないが、適中率でみる限り、平均的にみてロジットがやや良いが余り差異はない。分析手法上変数の選び易さは、①線型②判別③ロジット④教Ⅱ類の順であった。また、従属変数は、ロジットおよび線型モデルでは直接確率として解釈できるが、線型モデルでは0以下, 1以上となる場合が問題であり、判別および教Ⅱ類モデルでは解釈がしにくい。

以上の結果から、線型モデルで有力な変数の組を限定して、この限定組にロジットモデルを適用する方法が、一番有効であると考えられる。

【予測能力のチェック】 モデルの有効性は、現状再現性ではなく予測能力の良否で検討するべきであることは、明白である。ここでは、サンプルを2グループに分け、相互に、一方で算定したモデルで他方を予測

する方法をとった。予測能力は、これらの2ケースの決定係数 $R_c^2$ と適中率 $\alpha$ の平均値と比較することとした。変数の組は、個々の変数の安定性を重視し、定数項を含めて5変数のモデル、C [SPN, AVAILA, INCOME, RXXII] とD [SPN, INCOME, RXXII, DD2]とした。

表3: 変数組Cとサンプルサイズ  $n$  ( )内は $\tau$ 値

独立変数	サンプル-1 (N=57)	サンプル-2 (N=56)	サンプル-1+2 (N=113)
定数項	4.8026	6.1817	5.1706
AVAILA	-1.9757(2.63)	-1.8723(2.29)	-1.7433(3.37)
SPN	-0.8734(3.66)	-1.3302(3.03)	-0.9981(4.88)
INCOME	-0.3387(1.49)	-0.5115(2.00)	-0.4216(2.60)
RXXII	-1.7758(2.21)	-1.2907(1.82)	-1.5195(2.92)
適中率	89.47%	82.14%	85.84%
$\bar{P}^2$	.510	.568	.543
$R_c^2$	.593	.614	.607

表3は、変数組Cについて、サンプル数の変化に対するモデルの安定性を示すものである。サンプル数に対して変数の数を少なく押さえているため、また、サブグループ間で変数の分布に大きな変化がないため、サンプル数が半分になっても有効であることが認められる。変数組Dについても同様であるが、HOYUDAIを含む場合は、尤値が1以下になったり、計算が収束しないこととなった。この原因は、HOYUDAIの分布の偏り(2台以上は7名のみ)にあると考えられる。

表5の①と②、③と④の比較から、ロジットと線型では、知らず知らずのうちに予測能力はロジットの方が良いが、直中率 $\alpha$ については差がほとんどないことがわかる。

以上の分析より、このデータについては、現状再現性・予測能力ともに、ロジットモデルの方が線型モデルより良いが、大きな差はないことが明らかとなった。

## 2. モデルの拡張 ~ 重みの導入 ~

重みの導入は、代替物の代替性を考慮する試みである。交通手段として車を運転することも選択する場合に、他の代替手段(バスや電車)を利用することをどの程度意識しているかが、代替物の代替性である。代替性の高い代替物は、代替性の低い代替物よりも、モデルの算定において何らかの形で重視されるべきであり、より実際の選択状況を反映した精度の高いモデルが得られると考えられる。

実際に分析で用いた重みは、主に利用する交通手段と代替交通手段の利用頻度の差の程度(同じくらい、時々、めったに)であり、電車の車に対する相対の利用頻度 $<0.1 \sim 0.9>$ に修正してモデル化している。相対の利用頻度別度数分布を表4に示す。

表4: 相対の利用頻度と度数分布

交通手段	めったに	時々	同じくらい	計
電車 $n=1$	37 <0.90>	22 <0.70>	4 <0.50>	63
車 $P_{it}=0$	18 · 19	13 · 9	1 · 3	32 · 31
	26 <0.10>	12 <0.30>	12 <0.50>	50
	13 · 13	7 · 5	5 · 7	25 · 25

・ < > 内は、相対の利用頻度

・ 下段は、サブグループ1,2別の度数である

従属変数を、0.1~0.9 にできない(離散的な2項分布にあわない)ので、仮想的に各個人がすべて10回移動するとして、電車を9回、車を1回利用するという形で、係数の算定を行なうこととした。

重みづけの効果の判定基準としては、既述したサブ

グループ相互のモデルの予測能力の平均値を用いることとした。利用したモデルは、ロジットモデル、線型モデル、対数線型のロジットモデルの3モデルである。

表5: 予測能力の比較

番号	重みの有無	モデルタイプ	予測能力		
			$R^2$	$\alpha$	
①	無し	ロジット-C	.574	85.43%	・ モデルCの独立変数=定数項 SPN, AVAILA, RXX11, INCOME
②	無し	ロジット-D	.547	85.86%	
③	無し	線型-C	.527	83.93%	・ モデルDの独立変数=定数項 SPN, RXX11, INCOME, DD2
④	無し	線型-D	.514	86.72%	
⑤	有り	ロジット-C	.534	86.72%	・ 予測能力は、サブグループ相互の予測力、2グループの平均である。
⑥	有り	ロジット-D	.507	87.61%	
⑦	有り	線型-C	.526	85.83%	・ $R^2$ = 決定係数(自由度修正)
⑧	有り	線型-D	.482	86.73%	
⑨	有り	ロジット-C	.552	84.05%	・ $\alpha$ = 直中率
⑩	有り	(対数線型)D	.525	86.73%	

[重みの効果] 同じタイプのモデルで比較すると、

ロジットでも線型モデルでも①②と⑤⑥、③④と⑦⑧、 $R^2$ は重みのない方がよいが、 $\alpha$ では重みをもつ方がよい。また、重み付きのロジットモデルの場合、最尤推定法による場合(⑤⑥)と、対数変換により線型化して線型回帰を行なう場合(⑨⑩)とでは、最尤法による方が $R^2$ は低いが $\alpha$ は高い。

以上の結果、極めて少数サンプルによる比較であるが、ロジットモデルと線型モデルでは、ロジットモデルの方が、効用理論との関連が明快で、一般的に予測能力は高い。重みの効果は、このデータでは明瞭ではない。これは、重みの取り方とデータの特長性によると思われる。交通手段を選択する回数(トリップ発生量)の違いを考慮する等の改良を検討中である。

## 最後に、変数組C

について、線型モデルとロジットモデルでのパラメータの比較を行なう。表3と表6より、個々のパラメータの値(むしろ

表6. 変数組C: 線型モデルの結果

独立変数	サブグループ1(n=57)	サブグループ2(n=58)
定数項	1.5887	1.5019
AVAILA	-0.4548 (9.18)	-0.2156 (3.95)
SPN	-0.2249 (30.09)	-0.2258 (28.34)
INCOME	-0.593 (2.19)	-0.0591 (2.19)
RXX11	-0.3291 (8.35)	-0.3252 (5.01)
$R^2$	.545	.545
$\alpha$ (直中率)	87.72%	82.14%

< > 内はF値

相対比)が大きく変化しているのがわかる。この違いは重要であり、データによっては予測能力の差が生じることを示している。1つの方向として、より一般的なデータについて心理学的手法による要因の相対比を求めて、非集計モデルの分析結果と比較検討することが考えられる。

[参考文献] <1> T.A. Domencich, D McFadden, Urban Travel Demand. <2> 本稿その1, ロジットモデルの適用例, NewHolland, 1975