多項ロジットモデルによる非集計レベルの交通機関選択モデル

広島大学工学部　正会員　杉恵頼寧
東洋建設(株)　正会員　○田路隆茂

1．はじめに

交通機関選択モデルとして、非集計型ロジットモデルが、ここ数年注目をあび、短期交通政策の評価等に広く用いられるようになっている。しかし、その実用化にあたっては、まだいくつか検討すべき問題が残されている。そこで、本研究はその1つとして、交通機関を鉄道・バス　乗用車の3種に分類し、非集計レベルの交通機関選択をゾーンレベルに集計していく場合、どの集計手法が適しているかを比較検討する。次いで、非集計型ロジットモデルを作成する場合、必要となるサンプル数を決定するために、サンプル数を減少させると精度にどのような変化が表われるかを調べてみる。

2．モデルのキャリブレーション

非集計型ロジットモデルは、個人レベルのデータをそのまま用いて、個人の交通行動を解析しようとするものである。データには昭和54年9月に尾道・三原で行なわれた通勤アンケート票を用いた。多項ロジットモデルは次のように表わされる。

$$P_{(i,A_t)} = e^{Z_{it}} / \sum_{j \in A_t} e^{Z_{jt}} \qquad Z_{it} = \sum_{k=1}^{K} X_{itk} \cdot \beta_k$$

ただし、$P_{(i,A_t)}$：A_tからiの交通手段を選択する確率、A_t　tの人が利用可能な交通手段の集合、Z_{it}：tの人がiの交通手段を利用した場合の効用、X_{itk}：tの人に対するiの交通手段の変数k、β_k：変数kのパラメータ、K：変数の総数、

ここで用いる交通手段は鉄道・バス・乗用車の3種である。モデルのパラメータは最尤法によって求めた。それを表1に示しておく。それぞれのモデル式の検定としては、t値・χ^2値・ρ^2値及び的中率を用い、それぞれ、高い値を示しているほどそのモデル式は妥当であるといえる。しかし、総所要時間の変数を持ちいたモデルIとモデルIIIでは、併にパラメータの符号がプラスになっていて、一般に考えて不適である。そのため、総所要時間の変数を使用していないモデルIIを以下の研究に使用する。

表-1　モデルのパラメータ

変数＼モデル	I	II	III
性別	-1.732 (-5.84)	-1.770 (-6.33)	-1.363 (-5.11)
車保有	-5.219 (-8.90)	-5.277 (-9.08)	-4.398 (-8.46)
交通費 (100円)	-0.800 (-9.51)	-0.796 (-10.62)	
総所要時間 (10分)	0.302 (4.57)		0.137 (2.59)
アクセス距離 (km)	0.274 (3.54)		
エグレス距離 (km)	-0.572 (-4.74)	-0.537 (-4.60)	-0.264 (-2.89)
運行間隔 (10分)	-0.258 (-5.76)	-0.141 (-3.85)	-0.086 (-2.67)
乗換回数	-2.042 (-10.49)	-1.640 (-9.92)	-2.057 (-13.25)
的中率	88.62	86.75	83.91
ρ^2値	0.6372	0.6150	0.5111
χ^2値	1098	1058	881

（　）内はt値

3．非集計モデルの集計化

本研究では、非集計型ロジットモデルによる集計化の代表的手法である *Enumeration Method* と *Classification Method* を用いて、精度の比較検討を行なう。*Enumeration Method* と *Classification Method* の手法は次のようである。

- *Enumeration Method*　$R_{iN} = \sum_{t=1}^{n} P_{it} / n \times 100$ (%)

ただし、R_{iN}：Nゾーン着のトリップでiの交通機関分担率、
P_{it}：tの人がiの交通機関を選択する確率、
n：Nゾーン着のトリップ数、

- *Classification Method*　$R_{iN} = \sum_{j=1}^{m} \frac{n_j}{n} P_i(\bar{X}_j) \times 100$ (%)

ただし、n_j：サブグループjのトリップ数、m：サブグループの数、
$\bar{X}_j$：サブグループjにおける変数の平均値、
$P_i(\bar{X}_j)$：変数$\bar{X}_j$をモデル式に代入した時のi交通機関の選択確率、

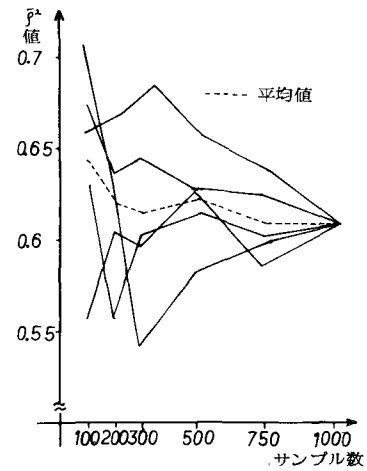
表-2　集計手法の精度比較

		分担率の推定		分担量の推定	
		RMS	%RMS	RMS	%RMS
Enumeration	乗車車	6.5	10.1	2.8	5.8
	バス	4.4	24.4	3.8	27.7
	鉄道	7.3	42.8	3.0	22.8
Classification I	乗車車	13.3	20.2	4.8	9.8
	バス	5.0	28.1	3.2	23.7
	鉄道	11.7	68.2	4.3	33.4
Classification II	乗車車	21.0	32.3	22.6	45.8
	バス	24.9	139.3	20.1	147.6
	鉄道	11.9	69.7	5.3	34.6

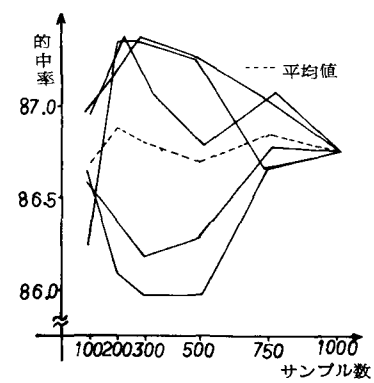
ここでサブグループとは、ゾーン内における性別、車保有・非保有者の組み合わせによる4つのサブグループのことである。アンケート調査には、個人ごとに各交通機関の指標について実績交通手段（実際に利用している交通手段）とその代替となる利用可能交通手段について答えた2種類のデータがある。よって各交通機関の指標（変数）の平均値$\bar{X}_j$を求める際にその両方を用いるのと、実績交通手段のみを使用する2通りがあり、前者を、*Classification* Ⅰとし、後者を*Classification* Ⅱとする。着ゾーンごとで10トリップ以上をもつ13ゾーンについて上記の3集計手法を用いて集計を行ない、交通機関別分担率を推定した。それと実績分担率との誤差をRMS、%RMS誤差で表わして（表2）各集計手法を比較検討してみる。分担量によるもの（分担率と総トリップ数の積の方が分担率によるものより誤差が少ない。これは、サンプル数の少ないゾーンはサンプル数の多いゾーンに比べて大きな誤差が出やすいが、分担率で表わすことによってサンプル数の多いゾーンと同じウェイトで扱われるためである。サンプル数の少ないゾーンでは、変数の平均値にかたよりが生じ、実績とかけはなれた分担率を推定する。*Classification* Ⅱの変数の平均値は実績交通手段しか用いないために、*Classification* Ⅰよりも平均値を求めるデータ数が少なくなり、*Classification* Ⅱの方がⅠよりも誤差が大きい。非集計レベルのモデルを個人レベルのままで使用し、それを集計した*Enumeration Method*の方が一般に*Classification Method*よりも優っている。これは*Classification Method*では変数の平均値を用いたことによる集計誤差が生じているためである。ただし、*Classification* Ⅰによる分担量の推定では、%RMS誤差が*Enumeration Method*と大差なく、集計化の手法として、今後利用出来そうである。

3　サンプル数の違いによる精度の検討

少数のデータでもって、精度の良いモデル式が得られるというのが、非集計型ロジットモデルの特徴の1つである。表1で求めたパラメータは、多項ロジットモデルを作成するのに有効である1019のデータを全て使用して求めたものである。それでは、そのサンプル数を減少させてモデル式を作成した場合、どのような精度を持つモデル式が得られるのかを調べてみる。1019の全データのうち、100, 200, 300, 500, 750$_{ss}$をランダムに取り出し、モデルのキャリブレーションを行なう。取り出し方を変えて以上の作業をそれぞれ5回繰り返した。その際に求まるρ^2値（モデル式に対するパラメータの説明力を表わす。）をプロットしたものが図1である。次に、そのようにして求めたモデル式でもって、全サンプルの利用交通機関を推定し、実績との的中率をそれぞれ求めた。それをプロットしたものが図2である。ρ^2値は1に近づくほど、的中率は高いほど、そのモデル式は妥当であるといえる。サンプル数の減少に伴なう精度の悪化という当初予測していた結果は見られないが、精度のばらつきが、サンプル数を減少させるに従がって大きくなっているのがわかる。的中率の差は、サンプル数750$_{ss}$の時、0.3%であったのが、500以下になると1.3%以上になっている。さらに的中率やρ^2値が良くても、モデル式のパラメータの符号が反対になるところがあり、データを半分以下にすることは問題が多い、パラメータの符号や、的中率、ρ^2値を検討してみて、データが750以上あれば全データで求めたモデル式と同じ程度の精度を期待しても良いと思われる。



図－1　サンプル数減少によるρ^2値の変化



図－2　サンプル数減少による的中率の変化