

広島大学工学部 正会員 杉恵頼等
 広島大学工学部 学生員 ○塚本冬文

1. はじめに

交通量の将来推計は、従来よりデータとゾーンレベルで扱う集計モデルが用いられている。これに対して、データと行動単位である個人レベルのデータを分析する非集計モデルの研究が最近盛んに行われている。これは変数として個人属性とモデルに組み込むことができるため、現実に近い行動特性が再現でき予測精度が高い。また集計モデルと比較して、調査データがはるかに少なくすむという利点と有している。しかし、実際の交通計画で必要なゾーンレベルでの分担率と求めらるためには、集計化の作業を必要とする。本研究は、二項選択ロジットモデルを用いて集計化における各集計手法の比較検討を行なったものである。

2. 調査の概要

本研究のデータは、広島市が昭和54年11月に実施した「交通問題に関する意識調査」の結果の内、通勤トリップに関するものから得た。調査の対象地域は、広島市と周辺6町からなる広島都市圏である。調査の対象者は総世帯数388,233の約21%にあたる8,229世帯のうち、通勤・買物目的に公共輸送機関あるいは自動車と利用する人として、そのうち、通勤目的の調査票はその該当者として5,590票配布し、5,083票回収した。データチェック後の有効票は4,647票であった。調査内容は①個人属性、②公共輸送機関のサービスの実態と満足度、③自動車利用の実態と満足度の3項目と基本とし、②と③については通常利用する交通機関と代替交通機関の両方についてとらえられよう構成されている。

3. モデルの作成

今回用いた二項選択ロジットモデルは次のとおりである。

$$P_j = 1 / (1 + e^{U_j}) \quad , \quad U_j = \beta_0 + \sum_{k=1}^m \beta_k (X_{ijk} - X_{j0k}) \quad \text{---(1)}$$

に於て、 P_j ：個人 j が公共輸送機関を選択する確率、 U_j ：個人 j の効用の差、 X_{ijk} 、 X_{j0k} ：公共輸送機関、自動車と利用する時の交通機関の選択要因 k に対する個人 j の値、 β_0 、 β_k ：パラメータ

パラメータ β_k ($k=0,1,\dots,m$)は最尤法により求めた。またモデルの作成にあたって、変数として組み込む調査項目すべてに解答している人とサンプルとして採用した。その結果全サンプル数2,395となった。

表1に使用した変数とその内容、および変数の平均値と土値を示す。土値は個々の変数と(1)式に代入したものの値である。これによると、年令、トリップ長、アクセス時間差、車内混雑度差の土値が小さく、交通機関の選択に対して説明力が小さいことがわかる。土値が大きいのは、性別、車の有無、通勤先、コスト差続いて総所要時間差、乗車時間差で交通機関の選択に説明力の大きい要因であるといえる。次に符号を検討してみると、総所要時間差から車内混雑度差までの変数は、変数の性質上パラメータの符号(土値の符号)はマイナスにならなければならないが、車内混雑度差を除いてすべてそれが満足している。

以上の結果をもとにパラメータの土値、重共線性、集計手法等を考慮

表1 使用変数の平均値と土値

変数	変数の内容	平均値(土値)
性別	男=1, 女=0	0.861 (-12.2)
年令(年)		37.5 (-1.7)
車の有無	保有=1, 非保有=0	0.854 (-78.0)
自宅の住所	旧広島市=1, 周辺部=0	0.198 (4.8)
通勤先	CBD=1, CBD以外=0	0.372 (-2.3)
トリップ長(km)	ゾーン間最短距離	11.1 (-2.3)
総所要時間差(分)	MT - CAR	19.8 (-10.6)
乗車時間差(分)	MT - CAR	4.9 (-9.4)
アクセス時間差(分)	MTアクセス時間	9.3 (-1.9)
エグレス時間差(分)	MT - CAR	4.6 (-3.5)
コスト差(1000円)	MT - CAR	-0.143 (-12.8)
乗換回数差(回)	MT乗換回数	0.440 (-4.9)
運行間隔差(分)	MT運行間隔	21.3 (-6.2)
定時性差	CARの遅れ(分)、所願=0 MTの遅れ=-1	-0.107 (-5.8)
車内混雑度差	(MTにおいて) すわれば1111, なければ0	0.503 (0.5)

(注)・CARは自動車、MTは公共輸送機関を示す。
 ・旧広島市とは、昭和46年当時の行政区域とす。

してモデルの作成を行なった結果が表2である。この場合、モデルの適合度を示す指標としては、的中率、 ρ 値、 σ 値を採用した。 ρ 値は回帰分析における重相関係数に相当する統計量で、1に近づくほど、また σ 値はその値が大きいほどモデルの適合度が高いことを示す。

4. 集計手法の比較検討

今回採用した集計手法は次のとおりである。

1) Enumeration Method

$$P_N = \sum_{i=1}^n P_i / n \quad (2)$$

ただし、 P_N ：グループN全体の選択確率、 P_i ：グループNに属する個人iの選択確率、 n ：グループNのサンプル数
これは個人ごとに推定した選択確率をグループごとに合計してグループの選択確率を求めらるものである。

2) Naive Method

$$P_N = P_i(\bar{x}_k) \quad (3)$$

ただし、 $\bar{x}_k$ ：各変数のグループ平均値
これは個人データによって測定されたモデルに、グループの平均値をインプットしてグループの選択確率を求めらるものである。

3) Classification Method

$$P_N = \sum_{c=1}^m n_c P_i(\bar{x}_{kc}) / N \quad (4)$$

ただし、 $\bar{x}_{kc}$ ：サブグループCにおける変数kの平均値、 n_c ：サブグループCのサンプル数、 N ：全サンプル数、 m ：サブグループ数
これはデータといくつかのサブグループに分け、サブグループごとに選択確率を求めサンプル数による重みづけを行なった後に、各々加え合わせて全体の選択確率を求めらるものである。

4) Taylor Series Method

$$P_N = P_i(\bar{U}_i) - \text{Var}(U_i) P_i(\bar{U}_i) (1 - P_i(\bar{U}_i)) (2 P_i(\bar{U}_i) - 1) / 2 \quad (5)$$

ただし、 $\bar{U}_i$ 、 $\text{Var}(U_i)$ ：効用の差 U_i の平均と分散

これは(1)式と $\bar{U}_i$ のまわりにTaylor展開して第3項までとったもので、Naive Methodに修正項を加えたものである。

集計化にあたっては、サンプル数と考慮してデータとCBD着のトリップに限り、旧広島市については大ゾーン分けとした。なお集計誤差と検討する指標としては% RMS誤差を採用した。

表3～5に各手法による集計結果を示す。これによると、予測精度の良いモデル5を用いたClassification, Enumeration, Taylor Seriesの集計精度が一番良い。これは、集計レベルをCBD着のゾーンペアに限定したが、モデルの変数「通勤先」がそれに直接関連しているためである。Naive, ClassificationとEnumerationとを比較すると後者の精度が良いが、これは前者に集計誤差が生じているためである。しかしその値はあまり大きくないので、変数の値がゾーンレベルでしかわからない将来予測には有効である。

Taylor Seriesは、効用の差の分散と求めなければならず将来予測には不適当であるが、精度が非常に良く短期予測に用いるとよい。

表2 集計化使用モデル (上層：パターナ

変数	モデル1	モデル2	モデル3	モデル4	モデル5
定数項	-0.101	0.595	1.095	1.296	0.763
性別	—	-0.832 (11.9)	—	-0.383 (4.9)	—
車の有無	—	—	-1.436 (17.2)	-1.289 (14.7)	-1.415 (16.9)
通勤先	—	—	—	—	0.603 (10.7)
乗車時間差	-0.011 (8.3)	-0.012 (8.6)	-0.012 (8.4)	-0.012 (8.5)	-0.012 (8.0)
アクセス時間差	-0.017 (3.6)	-0.016 (3.3)	-0.014 (2.7)	-0.014 (2.7)	-0.014 (2.6)
エgress時間差	-0.017 (4.3)	-0.019 (4.6)	-0.024 (5.5)	-0.024 (5.5)	-0.019 (4.3)
コスト差 (1000円)	-1.091 (12.0)	-1.150 (11.8)	-1.238 (11.1)	-1.260 (11.2)	-1.004 (8.9)
乗換回数差	-0.115 (2.5)	-0.088 (1.9)	-0.093 (1.9)	-0.080 (1.6)	-0.005 (1.1)
選行間隔差	-0.008 (6.1)	-0.009 (6.2)	-0.008 (5.3)	-0.008 (5.5)	-0.006 (3.7)
的中率	75.7	76.2	80.8	80.6	81.3
ρ 値	0.125	0.175	0.255	0.262	0.297
σ 値	383	532	772	793	889

表3 Naive, Classificationによる集計

手法	Naive	Classification			
変数項目	モデル1	モデル2 (性別)	モデル3 (車の有無)	モデル4 (乗車時間差)	モデル5 (通勤先)
平均値	0.377	0.394	0.395	0.397	0.517
平均残差	0.1342	0.1178	0.1170	0.1146	-0.0055
%RMS誤差	32.10	30.23	30.70	30.70	18.91

表4 Enumerationによる集計

変数項目	モデル1	モデル2	モデル3	モデル4	モデル5
平均値	0.386	0.398	0.404	0.408	0.514
平均残差	0.1251	0.4332	0.3621	0.3598	-0.0026
%RMS誤差	29.95	28.36	28.42	28.13	18.47

表5 Taylor Seriesによる集計

変数項目	モデル1	モデル2	モデル3	モデル4	モデル5
平均値	0.410	0.429	0.454	0.452	0.514
平均残差	0.1013	0.0829	0.0580	0.0598	-0.0025
%RMS誤差	27.28	24.27	23.11	23.23	17.63