

IV-189 選択モデルの比較に関する一考察

セントラルコンサルタント 正会員 ○細谷 涼子
鳥取大学工学部 正会員 奥山 育英
鳥取大学大学院 学生員 高梨 誠

1. はじめに

近年、交通計画における交通機関選択を中心として、ロジットモデルやそれを改良したモデルが多く利用されている。このロジットモデルは、以前から用いられていたプロビットモデルに代わるモデルとして開発されたものである。

そこで本研究では、選択モデルの中でも同一のデータ形式で解析方法のみが異なるプロビットモデルとロジットモデル、および判別分析について、その精度や特徴の考察を行うことを目的とする。

2. 本研究で取り扱う選択モデル

本研究では、前章で挙げた三つのモデルにおいて選択可能な選択肢が二つのみの場合を取り扱う。つまり、二つの選択肢 A と B があるとき、選択肢 A に対して効用 V_i を持つ i 番目のサンプルが選択肢 A を選択する確率 P_{Ai} は、プロビットモデルでは正規分布、ロジットモデルではワイブル分布を用いて

$$\begin{aligned} [\text{プロビット}] P_{Ai} &= \frac{1}{\sqrt{2\pi}} \int_{-V_i}^{\infty} \exp\left\{-\frac{z^2}{2}\right\} dz \\ [\text{ロジット}] P_{Ai} &= \frac{1}{1 + \exp\{-V_i\}} \end{aligned}$$

と表される場合である。ここで、効用 V_i は特性変数 x_{ki} ($k = 1, 2, \dots, K$) に対して線形であり

$$V_i = \sum_{k=1}^K \theta_k \cdot x_{ki} \quad (\theta_k : \text{未知のパラメータ})$$

とする。但し、 $\theta_1 \cdot x_{1i}$ は定数項であり、それは常に $x_{1i} = 1$ である選択肢固有ダミー変数である。

まずプロビットモデル分析とロジットモデル分析では、実際の選択結果と選択確率を用いて、最尤法によって効用関数 $V(\vec{\theta})$ を推定する。

また判別分析では、実際に選択肢 A を選択したグループと選択肢 B を選択したグループの効用 V_i ができるだけ分かれるように、つまり相関比が最大と

なるように効用関数を推定する。

3. 選択モデルの比較の方法

昌中の研究¹⁾にあるデータ(データ A)と Daganzo の研究²⁾にあるデータ(データ B)を用いて、以下のような手順で選択モデルの比較を行う。

- (1) 三つの分析方法によって、効用関数 $V(\vec{\theta})$ を推定し、プロビット、ロジットに関しては対数尤度も求める。但し、データ B は選択肢固有ダミー変数以外の特性変数が一つのため、判別分析は行えないので相関比のみを求める。
- (2) 相関比 P^2 を求める。相関比は選択結果ごとの効用 V_i の分離度を示す指標であり、1に近いほどグループ間の分離が明確であることを示す。
- (3) しきい値を用いた的中率 R_S を求める。つまり、得られた効用のうちのある値をしきい値として、各サンプルの選択結果を推定する。この推定と実際の選択結果との一致の程度を示す変数を S_i として

$$S_i = \begin{cases} 1 & (\text{推定と結果が一致}) \\ 0 & (\text{推定と結果が不一致}) \end{cases}$$

とすると、的中率 R_S は

$$R_S = \frac{1}{N} \sum_{i=1}^N S_i$$

と定義することができる。

- (4) 確率を用いた的中率 R_R を求める。ここでは、同じような効用をとるサンプルを一つのグループとして、全サンプルを M 個のグループに分ける。次に各グループの選択肢 A を選択するサンプル数 T_{Am} を推定する。そして、実際に選択肢 A を選択したサンプル数を N_{Am} とする

と、的中率 R_R は次式のように表される。

$$R_R = \frac{N - H}{N}$$

$$H = \sum_{m=1}^M |N_{Am} - T_{Am}|$$

4. 計算結果と考察

計算結果を以下の表-1、表-2 に示す。なお、データ B についての相関比は 0.4661 であった。

表-1 データ A の分析結果

パラメータ	プロビット	ロジット	判別
θ_1	0.4055 (28.0)	0.7363 (27.0)	—
θ_2	1.350 (93.1)	2.561 (93.8)	-8.731 (17.4)
θ_3	-0.0145 (-1)	-0.0273 (-1)	0.501 (-1)
θ_4	0.0994 (6.9)	0.1940 (7.1)	-3.218 (6.4)
L	-56.870	-55.860	—
P^2	0.2284	0.2286	0.3238
R_S	0.9488	0.9458	0.9337
R_R	0.9508	0.9529	—

() 内の数値は $\theta_3 = -1$ としたときの
パラメータ間の比である。

表-2 データ B の分析結果

パラメータ	プロビット	ロジット
θ_1	-2.282 (-6.3)	-4.110 (-6.2)
θ_2	0.3626 (1)	0.664 (1)
L	-103.213	-103.629
P^2	0.3220	0.3220
R_S	0.8991	0.8991
R_R	0.9518	0.9427

() 内の数値は $\theta_2 = -1$ としたときの
パラメータ間の比である。

これらの結果より分かることは、パラメータ推定値自体は扱うモデルによってかなり異なるが、それぞれのモデルにおいてパラメータ間の比を計算すると、プロビットとロジットに関しては同じような値

が得られた。また、対数尤度、相関比、的中率といったモデルの精度を表す各指標も、三つのモデル間では差がほとんどみられなかった。

これは、プロビットモデルとロジットモデルでは効用関数の誤差項の分布形が違うだけであり、その分布形も類似していて、なおかつ分析方法も変わらないためであると考えられる。

ところで、判別分析に関しては、その解析方法が他の二つのモデルと全く異なるにもかかわらず、結果として同程度の的中率を示していたことから、効用関数の誤差項を考慮しない判別分析もそれを用いるときの条件によっては有効であるといえる。

5. おわりに

二種のデータを用いてプロビットモデルとロジットモデルを比較した結果、精度にはほとんど差がみられなかった。従って、両者を比べた場合、二項選択モデルではどちらも容易に扱えるが、多項選択モデルになるとプロビットモデルの扱いが複雑になることを考慮すれば、一般的にプロビットよりロジットを用いる方が妥当であるといえる。

また、確率を用いた的中率 R_R は、しきい値を用いた的中率 R_S より向上するため、特に同じ効用値をとるサンプルの数がある程度以上のデータでは、モデルの精度をみるための指標として、あるいは現象予測としてこのような方法を利用すると、より現実に即した判断や予測が可能になると考えられる。

このことは判別分析についても同様であり、本来判別分析を用いた的中率の計算や予測においては確率を考えないが、サンプル数の多いデータでは、各選択肢ごとの効用の分布から効用に対する選択確率を求め、プロビットモデルやロジットモデルのようの中率や予測に適用し得るであろう。

参考文献

- 1) 畠中・喜多：高速道路流入部におけるギャップアクセプタンス挙動の解析と加速車線長の決定法，第42回土木学会中国四国支部研究発表会講演概要集，pp.432-433，1990。
- 2) Carlos F. Daganzo: Estimation of Gap Acceptance Parameters within and across the Population from Direct Roadside Observation, Transportation Research, vol.15B, pp.1-15, 1981.