

共分散構造分析を用いたプロビット回帰モデル

立教大学大学院 齊藤 俊則 SAITO Toshinori

I はじめに

本研究は、共分散構造分析を用い、消費者の選択行動を記述・予測するモデルを作成するものである。柳井・築掛・前川・市川(1997)によれば、カテゴリカル因子分析と項目反応理論は同一のものを見なせ、また豊田(1997ab)によれば、項目反応理論の考え方を共分散構造分析に導入することにより、離散型の基準変数に対する確率予測が可能であることが示されている。後述するこの方法を選択確率の予測に応用し、あわせてその方法の特徴について紹介する。またここでの分析の対象は国内旅行先選択である。

本研究では、確率的選択モデルに対する潜在変数の導入の試みと個人差異質性の取り扱いの1つの方法として共分散構造分析を用いた選択モデルを紹介する。

II 項目反応理論とプロビット回帰分析

まず、一般に多変量解析における離散変数の取り扱い、順序尺度データについては多分・多双列相関の考え方による対応が可能である(豊田 1997ab、狩野 1997)。但し、名義尺度データの場合はこの限りではない。

次に、項目反応理論(item response theory)とは、学力試験や心理検査から得られるテストデータを取り扱う統計モデルで、測定される特性を潜在変数によって表現し、学力を推定したり、テスト項目の特徴を記述するものである(芝 1991、池田 1994、豊田 1997ab)。

ここでは説明を簡潔にするために、2つのカテゴリを持つ順序変数1つと潜在変数を1つもモデルを考え、このモデルのパラメータ推定について説明する。まず、このモデルは、

$$z_i = \alpha f + e_i \quad (1)$$

と表現できる。ここで、 z_i は観測変数(順序変数)・ α は因子負荷量・ f は潜在変数(因子)・ e_i は誤差変数である。この z の背後に観測されない値 u という変数を仮

が観測されていると考えられる。この τ を閾値という。 $f = \theta$ という特定の値をとる下位の母集団から得た観測対象が $u_i = 1$ となる確率は、平均 $\alpha_i \theta$ 、分散 $1 - \alpha_i^2$ の正規分布の密度関数の定積分として

$$p(u_i = 1|\theta) = \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi(1-\alpha^2)}} \exp\left[-\frac{1}{2} \frac{(z_i - \alpha\theta)^2}{1-\alpha^2}\right] dz_i \quad (2)$$

と表現される。ここで、

$$h = \phi(z_i) = z_i - \alpha\theta / \sqrt{1-\alpha^2} \quad (3)$$

とし、 h を代入すると

$$p(u_i = 1|\theta) = \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{h^2}{2}\right] dh \quad (4)$$

式(4)は原点を中心として対象な関数だから、

$$p(u_i = 1|\theta) = \int_{-\phi(\tau_{ij})}^{\phi(\tau_{ij})} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{h^2}{2}\right] dh \quad (5)$$

となる。積分区間は、マイナス無限大から $-\phi(\tau_{ij})$ までであるが、この $-\phi(\tau_{ij})$ は、

$$-\phi(\tau_{ij}) = \frac{\alpha\theta - \tau_{ij}}{\sqrt{1-\alpha^2}} \quad (6)$$

であり、仮に

$$a_i = \alpha / \sqrt{1-\alpha^2} \quad (7)$$

$$b_i = \tau_{ij} / \alpha \quad (8)$$

とすると、式(5)は以下のようになる。

$$p(u_i = 1|\theta) = \int_{-a_i b_i}^{a_i b_i} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{h^2}{2}\right] dh \quad (9)$$

なお、式(14)の b_i を「困難度」といい、 a_i を「識別力」という。 b_i ：「困難度」は、求まる確率が50%の z の値を表わし、 a_i ：「識別力」は、項目特性曲線の傾きを表わす。

ここでプロビット回帰分析には、以下のように

$$Z_i = A_j f + \varepsilon_i \quad (11)$$

とする。Zは観測変数（順序変数）・Aは因子荷
量・ ε は潜在変数（因子）・ e_i は誤差変数である。
となる。ここで、式(10)を考慮して式(5)積分区間を変
更するが良い。

3. アプローチの方法

本研究では、前節で説明したプロビット回帰分析に
より、旅行先選択の選択率の予測を行う。ここで、旅
行というもの是一般の消費者にとっては、日雑品購入
の場合とは異なり、高頻度での購入を行うものでは
なく、年に数回しか消費（購入）できないものであ
ろう。この観光消費の特徴を考慮して、まず、支出（価
格・費用）に関する情報をモデルに採用する。さらに、
これらの2つの情報については1回の旅行に対する規
定としてではなく、年間の合計として扱った。という
のも旅行というものは、相対的には計画性の高い消費
（商品）であると考えられることと、家計経済から考
えても週・月単位の自由裁量金から消費できる性質の
ものではないと考えられるからである。つまり、旅行
支出と年間で旅行にさける余暇時間の2変数が旅行先
選択への制約条件とし、その選択行動に影響を与えて
いると思われる。また、国内旅行の場合居住地区の影
響が大きいと考えられるために、関東圏在住者の子
タのみを用いることにする。

III 実証研究

上記の仮説に基づき以下のようなモデルを作成し分
析を行った。結果は以下の通りである。まず、分析の
結果、最も適合度の高かったモデルを以下に示す。

図1

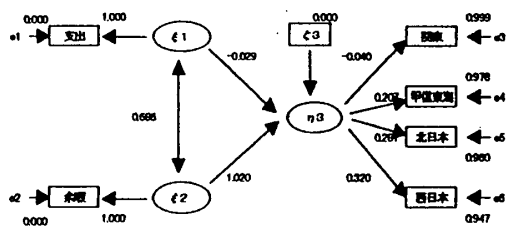


図1のモデルから導いた選択確率の推定は、以下の
通りであった。

表2

	実測値%	推定値%
関東	51.0	50.5
甲信越・東海	60.1	64.0
北日本	27.6	29.4
西日本	20.4	21.7

また、各旅行先が選択への困難度と識別力は以下の
通りであった。

表3

	困難度	識別力
関東	0.638	-0.040
甲信越・東海	-1.243	0.212
北日本	2.112	0.293
西日本	2.587	0.338

本研究では、上記モデルを近畿・中部在住者の
データにも適用し、算出した「困難度」の比較により、
観光地間の競争構造を把握する。また、ロジットモデ
ルの持つ「I.I.A.」「補償型ルール」「個人の同質
性」の問題点への対応についても論証する。

芝 祐順 (1991) 「項目反応理論: 基礎と応用」
東京大学出版会
池田 央 (1994) 「現代テスト理論」朝倉出版
豊田 秀樹 (1997 a) 立教大学大学院・東京大学大
学院 講義資料
豊田 秀樹 (1997 b) 「共分散構造分析による応用
研究」資料 心理測定研究所 第4回心理測定研究所
セミナー
柳井 晴夫・繁樹 算男・前川 眞一・市川 雅教 (1997)
「因子分析」朝倉書店
狩野 祐 (1997) 「AMOS・EQS・LISRELによるグ
ラフィカル多変量解析」