

エントロピー理論を用いた離散型選択モデルの導出と
パラメータ推定

A Derivation of Disaggregate Behavioral Model by Entropy Theory
and Estimation of Parameter

2005年 2月

坂野 匡弘

Masahiro Sakano

エントロピー理論を用いた離散型選択モデルの導出と パラメータ推定

A Derivation of Disaggregate Behavioral Model by Entropy Theory and Estimation of Parameter

坂野 匡弘

Masahiro Sakano

Abstract

Disaggregate behavioral model has been applied for transportation mode. This study clarified problems in the estimation process of its parameter. The distribution of the error term of the model, an average and dispersion are obvious and the maximum likelihood method or the nonlinear least square method may do the parameter estimation of this model. In this study, from the viewpoint of transportation mode choice as an alternative action, model derivation by the Entropy theory that Shannon devised was carried out. The distribution of the error term became obvious by a model's following multinomial distribution, too. The nature of the error term of the model which became distinct was used, and parameter estimation was done by maximum likelihood method.

The value requested, however, Gauss-Newton method that the second differential calculus omitted was used for the calculation. That made likelihood the biggest, since it is close to the linear form, though a model is a nonlinear form.

要約

現在、交通機関選択モデルとして多用されている離散型選択モデルを取り上げ、パラメータ推定プロセスにおける問題点を明らかにした。モデルパラメータを非線形最尤法または非線形最小二乗法で行うために誤差項の分布が明確でなくてはならない。本研究では交通選択が二者択一の行動であることに着目し、Shannon が考案したEntropy理論によるモデル導出を行うことによりモデルの分布を明確にし、非線形最尤法によりパラメータの推定を行った。非線形最尤法では2次の微分を省略したガウス・ニュートン法を用いたところ、ニュートン法の結果とはほぼ一致した。

2004年9月29日 受理

はじめに

交通行動分析を集団の特性として分析した集団分析から、個人の行動に重点をおいた離散型選択モデルが開発され、多くの分野で利用されている。特に交通計画の分野においては交通機関選択に欠かせない重要なモデルとなっている。個人の交通機関の選択は、ある交通機関を選択するかしないかの二者択一情報として捉えることができる。このことは情報の1ビットに相当し、選択確率は0、1の値をとる。このことに着目しShannonが考案したEntropyを利用し離散型選択モデルの導出を行った。交通機関を選択する確率分布は多項分布を仮定してモデルを誘導したので誤差項の分布、さらに分布の平均と分散も明確になった。

パラメータ推定はモデルの誤差項の分布、分散が明確になったのでAitkenの一般化最小二乗法または最尤法の適用が可能となった。双方の結果ととして得られる正規方程式は同じになるが本研究では最尤法を用いた。正規方程式はパラメータに関して非線形となるが比較的線形に近いことからガウス・ニュートン法を用いたが良好な結果が得られた。モデル適用例として北海道と本州間の生活雑貨の輸送データを用いた。北海道、本州間の全輸送の8割が内航運輸に委ねられ残りの2割の生活雑貨がJRとフェリーでシェア争いにしのぎを削っている。荷主がJRとフェリーを選択する状況を離散型選択モデルを用いて数値計算を行ったが高い的中率が得られた。

1 エントロピー理論を用いた離散型選択モデルの構築

(1) エントロピーと選択行動

個人 t の交通機関の選択は、ある交通機関を選択するか、しないかの二者択一の情報として捉えることができる。このことは情報理論における1ビットの情報量に相当する。

いま m 個の交通機関から交通機関 i を選択する、またはしないという事象の情報量はShannonによると

$$w_{it} = -\log p_{it} \quad (i=1,2,\dots,m;t=1,2,\dots,n) \quad (1.1)$$

ただし、 p_{it} : 個人 t が交通機関 i を選択する確率

となる。交通機関を選択するという情報量の期待値は

$$\begin{aligned} E(w_t) &= p_{1t}w_{1t} + p_{2t}w_{2t} + \dots + p_{mt}w_{mt} \\ &= -p_{1t} \log p_{1t} - p_{2t} \log p_{2t} - \dots - p_{mt} \log p_{mt} \\ &= -\sum_{i=1}^m p_{it} \log p_{it} \end{aligned} \quad (1.2)$$

となる。これがShannonのエントロピー (Entropy) であり、不確定な状況を確定するのに要する平均情報量となる。つぎに、現在良く利用されている多肢選択の離散型選択モデル(ロジットモデル) をエントロピーを用いて導出する。

交通機関 i を選択する効用 z_{it} を要因 x_{ijt} ($j=1,2,\dots,k$) の一次結合で表すと

$$z_{it} = \beta_0 x_{it1} + \beta_1 x_{it2} + \dots + \beta_k x_{itk} \quad (i=1,2,\dots,m) \quad k: \text{要因の数} \quad (1.3)$$

となる。

事象 w_{it} ($i=1,2,\dots,m$) に対する選択確率を $p_{1t}, p_{2t}, \dots, p_{mt}$ とすると情報の期待値は

$$E(w_t) = p_{1t} w_{1t} + p_{2t} w_{2t} + \dots + p_{mt} w_{mt} \quad (1.4)$$

となる。

(2) 離散型選択モデルの導出

離散型選択モデルを導出するためにエントロピー H を最小にするような確率 $p_{1t}, p_{2t}, \dots, p_{mt}$ を求める。しかし、条件1のみでラグランジュ乗数法でエントロピーを最小にすると $p_{1t} = p_{2t} = \dots = p_{mt} = 1/m$ という不確定な状況が得られる。本論文では $p_{1t}, p_{2t}, \dots, p_{mt}$ のどれかが1で他は0となる状況を求めるのが目的である。確定状況を得るためにエントロピー H が $H \geq 0$ の範囲で最小になるようにするため、次の条件2を付加した。

i) 目的関数

エントロピー H を最小にする。

$$H = -\sum_{i=1}^m p_{it} \log p_{it} \Rightarrow \text{最小} \quad (1.5)$$

ii) 条件1

確率 p_{it} ($i=1,2,\dots,m$) の和は1である。

$$\sum_{i=1}^m p_{it} = 1 \quad (t=1,2,\dots,n) \quad (1.6)$$

iii) 条件2

$$\sum_{t=1}^n \sum_{i=1}^m p_{it} x_{ijt} = \sum_{t=1}^n \sum_{i=1}^m \delta_{it} x_{ijt} = \bar{z}_j \quad (j=1,2,\dots,k) \quad (1.7)$$

δ_{it} : 事象 w_{it} に対応する実現確率、 x_{ijt} : 効用 z_{it} の要因

ここで未定乗数として $\lambda_0, \lambda_1, \dots, \lambda_k$ を用い、Lagrange関数を

$$F(p_1, \dots, p_m, \lambda_0, \dots, \lambda_k) = -\sum_{t=1}^n \sum_{i=1}^m p_{it} \log p_{it} + \lambda_0 \left(\sum_{i=1}^m p_{it} - 1 \right) + \sum_{j=1}^k \lambda_j \left(\sum_{t=1}^n \sum_{i=1}^m p_{it} x_{ijt} - \bar{z}_j \right) \quad (1.8)$$

と設定する。

$F(p_1, p_2, \dots, p_m, \lambda_0, \lambda_1, \dots, \lambda_k)$ を最大にする p_{it} を求めるために

$$\frac{\partial F}{\partial p_i} = 0 \quad (i = 1, 2, \dots, m) \quad , \quad \frac{\partial F}{\partial \lambda_j} = 0 \quad (j = 0, 1, 2, \dots, k) \quad (1.9)$$

と微分をおこなう。(1.9) 式より

$$\frac{\partial F}{\partial p_i} = -\log p_{it} - 1 + \lambda_0 + \sum_{j=1}^k \lambda_j x_{itj} = 0 \quad (i = 1, 2, \dots, m) \quad (1.10)$$

$$\frac{\partial F}{\partial \lambda_0} = \sum_{i=1}^m p_{it} - 1 = 0 \quad (1.11)$$

が得られる。(1.10) 式より

$$p_{it} = \exp(\lambda_0 - 1) \exp\left(\sum_{j=1}^k \lambda_j x_{itj}\right) \quad (1.12)$$

が導かれる。この (1.12) を (1.11) 式に代入すると

$$\begin{aligned} \sum_{i=1}^m p_{it} &= \sum_{i=1}^m \exp(\lambda_0 - 1) \exp\left(\sum_{j=1}^k \lambda_j x_{itj}\right) = 1 \\ \exp(\lambda_0 - 1) \sum_{i=1}^m \exp\left(\sum_{j=1}^k \lambda_j x_{itj}\right) &= 1 \end{aligned} \quad (1.13)$$

となる。

(1.12)、(1.13) 式より

$$p_{it} = \frac{\exp\left(\lambda_0 + \sum_{j=1}^k \lambda_j x_{itj}\right)}{\sum_{i=1}^m \exp\left(\lambda_0 + \sum_{j=1}^k \lambda_j x_{itj}\right)} \quad (1.14)$$

が導びかれた。(1.14)式の指数部は効用 z_{it} を表すので

$$z_{it} = \lambda_0 + \lambda_1 x_{it1} + \lambda_2 x_{it2} + \dots + \lambda_k x_{itk} \quad (t = 1, 2, \dots, n : i = 1, 2, \dots, m) \quad (1.15)$$

と表すことができる。したがって一般性を失うことなく、(1.15) 式を次式のように置く。

$$\beta' x_{it} = \beta_0 x_{it0} + \beta_1 x_{it1} + \dots + \beta_K x_{itK} \quad (1.16)$$

ここに $x_{i0} = 1$ とすると β_0 は定数項を表すことになる。

よって導出された離散型選択モデルは

$$p_{ii} = \frac{\exp(\beta' x_{ii})}{\exp(\beta' x_{i1}) + \exp(\beta' x_{i2}) + \cdots + \exp(\beta' x_{im})} \quad (i = 1, 2, \dots, m) \quad (1.17)$$

$$\beta' x_{ii} = \beta_0 x_{i0} + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$$

ここに

p_{ii} : 交通機関 i の選択確率

β_j : モデルのパラメータ

k : 交通機関 i の選択の効用の要因数

である。

2 最尤法による離散型選択モデルの推定

1 で導出した離散型選択モデルのパラメータ推定に先立ち (1.17) 式を次のように拡張する。

$$p_{ii} = \frac{\exp(\beta' x_{ii})}{\exp(\beta' x_{i1}) + \exp(\beta' x_{i2}) + \cdots + \exp(\beta' x_{im})} + u_{ii} \quad \begin{cases} (t = 1, 2, \dots, n) \\ (i = 1, 2, \dots, m) \end{cases} \quad (2.1)$$

$$\beta' x_{ii} = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}$$

ここに、

u_{ii} : 誤差項, p_{ii} : 選択確率, $\beta' x_{ii}$: 効用関数で $x_{i1} = 1$, β : 効用関数の係数

k : 効用の要因の数, t : 交通機関を選択する個人 n : 個人の数,

m : 選択する交通機関の数

である。

(2.1) 式において p_{ii} は多項分布で期待値と分散は

$$E(p_{ii}) = \sum_{i=1}^n p_{ii} \quad V(p_{ii}) = \sum_{i=1}^n p_{ii}(1 - p_{ii}) \quad (2.2)$$

十分大なる n に対して p_i は中心極限値定理により $N\left(\sum_{i=1}^n p_{ii}, \sum_{i=1}^n p_{ii}(1 - p_{ii})\right)$ の正規分布に近づく。

(2.2) 式より誤差項の分散

$$V(p_{it}) = V\left(\frac{\exp(\beta' x_{it})}{\exp(\beta' x_{1t}) + \exp(\beta' x_{2t}) + \cdots + \exp(\beta' x_{mt})} + u_{it}\right) = V(u_{it}) = \sum_{t=1}^n p_{it}(1 - p_{it}) \quad (2.3)$$

となる。ベクトル p_i, u_i を

$$p_i = \begin{bmatrix} p_{i1} \\ p_{i2} \\ \vdots \\ p_{in} \end{bmatrix}, \quad u_i = \begin{bmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{in} \end{bmatrix} \quad (i = 1, 2, \dots, m) \quad (2.4)$$

$$p' = [p_1, p_2, \dots, p_m], \quad u' = [u_1, u_2, \dots, u_m] \quad (2.5)$$

として、誤差項 u_i に関し次の仮定をおく

〔仮定 1〕 誤差項の期待値は 0 である。

$$E(u_{it}) = 0 \quad (i = 1, 2, \dots, m)$$

〔仮定 2〕 誤差項の分散は

$$V(u_{it}) = p_{it}(1 - p_{it}) \quad (t = 1, 2, \dots, n)$$

で共分散は 0 したがって分散共分散行列は

$$\Sigma_i = \begin{bmatrix} p_{i1}(1 - p_{i1}) & 0 & \cdots & \cdots & 0 \\ 0 & p_{i2}(1 - p_{i2}) & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & p_{in}(1 - p_{in}) \end{bmatrix} \quad (i = 1, 2, \dots, m)$$

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & 0 & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma_m \end{bmatrix}$$

とする。すなわち分散は観測点毎に不均一とし、共分散は 0 とする。

〔仮定 3〕 誤差項 u_i は

$$u_{it} \sim N\left(0, \sum_{t=1}^n p_{it}(1 - p_{it})\right)$$

とする。

誤差項 $u_{it} (i=1,2,\dots,m; t=1,2,\dots,n)$ は互いに独立であるとする。このとき u_{it} の同時分布密度は次のようになる。

$$\begin{aligned} p(u_{11}, u_{12}, \dots, u_{1n}, \dots, u_{m1}, u_{m2}, \dots, u_{mn}) &= \prod_{i=1}^m \prod_{t=1}^n p(u_{it}) \\ &= p(u_{11})p(u_{12}) \cdots p(u_{1n}) \cdots p(u_{m1})p(u_{m2}) \cdots p(u_{mn}) \end{aligned} \quad (2.6)$$

ここで Σ_i を誤差項 u_i の分散共分散行列とすると、 $p(u_{it})$ は u_{it} の密度関数は

$$f(u_i) = (2\pi)^{-n/2} |\Sigma_i|^{-1} \exp\left(-\frac{1}{2} u_i' \Sigma_i^{-1} u_i\right) \quad (i=1,2,\dots,m) \quad (2.7)$$

となる。また、

$$p(u_i) = f(u_i) \quad (i=1,2,\dots,m)$$

であるので (2.6) 式の関係は次式によって表すことができる。

$$\begin{aligned} f(u_{i1}, u_{i2}, \dots, u_{in}) &= \prod_{t=1}^n f(u_{it}) \\ &= (2\pi)^{-\frac{n}{2}} |\Sigma_i|^{-n} \exp\left(-\frac{1}{2} u_i' \Sigma_i^{-1} u_i\right) \quad (i=1,2,\dots,m) \end{aligned} \quad (2.8)$$

ここで

$$u' = [u_1, u_2, \dots, u_m] \quad , \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma_m \end{bmatrix} \quad (2.9)$$

とし、標本観察値の尤度を L とすると、(2.8) 式より

$$L = (2\pi)^{-\frac{n}{2}} |\Sigma|^{-n} \exp\left(-\frac{1}{2} u' \Sigma^{-1} u\right) \quad (2.10)$$

となる。両辺の対数をとると

$$L^* = \log L = -\frac{n}{2} \log(2\pi) - n \log |\Sigma| - \frac{1}{2} u' \Sigma^{-1} u \quad (2.11)$$

と表すことができる。ここで p_{it} の推定値を $\hat{p}_{it}$ 、 β の推定値を $\hat{\beta}$ とすると

$$\hat{p}_{it} = \frac{\exp(\hat{\beta}' x_{it})}{\exp(\hat{\beta}' x_{i1}) + \exp(\hat{\beta}' x_{i2}) + \cdots + \exp(\hat{\beta}' x_{im})} \quad \begin{cases} (t=1,2,\dots,n) \\ (i=1,2,\dots,m) \end{cases} \quad (2.12)$$

$$\hat{\beta}' x_{it} = \hat{\beta}_1 x_{it1} + \hat{\beta}_2 x_{it2} + \cdots + \hat{\beta}_k x_{itk}$$

と表すことができるので、 p_{it} と $\hat{p}_{it}$ との残差ベクトルを改めて $\hat{f}_{it}(\hat{\beta})$ とおくと

$$\begin{aligned}\hat{f}_{it}(\hat{\beta}) &= p_{it} - \hat{p}_{it} \\ &= p_{it} - \frac{\exp(\hat{\beta}'x_{it})}{\exp(\hat{\beta}'x_{1t}) + \exp(\hat{\beta}'x_{2t}) + \cdots + \exp(\hat{\beta}'x_{mt})}\end{aligned}\quad (2.13)$$

となる。さらに、(2.11) 式の対数尤度は

$$L^* = \log L = -\frac{n}{2} \log(2\pi) - n \log |\hat{\Sigma}| - \frac{1}{2} \hat{f}_i(\hat{\beta})' \hat{\Sigma}^{-1} \hat{f}_i(\hat{\beta}) \quad (2.14)$$

となる。ここで $f_i(\hat{\beta}) \in C^2$ とし、 $f_i(\hat{\beta})$ のヤコビ行列 (Jacobian Matrix) を

$$J'_i(\hat{\beta}) = \nabla f_i(\hat{\beta}) = \begin{bmatrix} \frac{\partial \hat{f}_{i1}}{\partial \hat{\beta}_1} & \frac{\partial \hat{f}_{i1}}{\partial \hat{\beta}_2} & \cdots & \frac{\partial \hat{f}_{i1}}{\partial \hat{\beta}_k} \\ \frac{\partial \hat{f}_{i2}}{\partial \hat{\beta}_1} & \frac{\partial \hat{f}_{i2}}{\partial \hat{\beta}_2} & \cdots & \frac{\partial \hat{f}_{i2}}{\partial \hat{\beta}_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{f}_{im}}{\partial \hat{\beta}_1} & \frac{\partial \hat{f}_{im}}{\partial \hat{\beta}_2} & \cdots & \frac{\partial \hat{f}_{im}}{\partial \hat{\beta}_k} \end{bmatrix} \quad (i=1,2,\dots,m) \quad (2.15)$$

$$J(\hat{\beta}) = [J_1(\hat{\beta}), J_2(\hat{\beta}), \dots, J_m(\hat{\beta})] \quad (2.16)$$

とし、 $\nabla = (\partial/\partial \hat{\beta}_1, \partial/\partial \hat{\beta}_2, \dots, \partial/\partial \hat{\beta}_k)$ を導入すると (2.14) 式を最大にする解は、

$$\nabla L^*(\hat{\beta}) = J'(\hat{\beta}) \hat{\Sigma}^{-1} f(\hat{\beta}) = 0 \quad \nabla L^*(\hat{\beta}) \in R^{k \times k} \quad (2.17)$$

の極値を求めることになり、非線形連立方程式を解く問題に帰着される。

この方程式を次のニュートン法で解く。 ∇L^* を $\hat{\beta}^l$ の近傍で 2 次の項まで Taylor 展開する

$$\nabla L^*(\hat{\beta}) = \nabla L^*(\hat{\beta}^l) + \nabla^2 L^*(\hat{\beta}^l)(\hat{\beta} - \hat{\beta}^l) + \frac{1}{2!}(\hat{\beta} - \hat{\beta}^l)' \nabla^3 L^*(\hat{\beta}^l)(\hat{\beta} - \hat{\beta}^l) + \cdots \quad (2.18)$$

である。 $\nabla^2 L^*(\hat{\beta}^l)$ が正定値であるとする (2.17) 式の 2 次関数の極値は最小点となるから (2.18) 式の 2 次の項で打ち切った式を 0 とおくと

$$\nabla^l L^*(\hat{\beta}^l) + \nabla^2 L^*(\hat{\beta}^l)(\hat{\beta} - \hat{\beta}^l) = 0 \quad (2.19)$$

この (2.19) 式の解を次の点 $\hat{\beta}^{l+1}$ とする。すなわち $\nabla^2 L^*(\hat{\beta}^l)$ が正則ならば

$$\hat{\beta}^{l+1} = \hat{\beta}^l - \{\nabla^2 L^*(\hat{\beta}^l)\}^{-1} \nabla^l L^*(\hat{\beta}^l) \quad (2.20)$$

として与えられる。 $\nabla^2 L^*(\hat{\beta}^l)$ は $\hat{\beta}^l$ におけるヘッセ行列 (Hessian) で

$$\nabla^2 L^*(\hat{\beta}^l) = \begin{bmatrix} \frac{\partial^2 L^*}{\partial \hat{\beta}_1^{l^2}} & \frac{\partial^2 L^*}{\partial \hat{\beta}_1^l \partial \hat{\beta}_2^l} & \cdots & \frac{\partial^2 L^*}{\partial \hat{\beta}_1^l \partial \hat{\beta}_k^l} \\ \frac{\partial^2 L^*}{\partial \hat{\beta}_2^l \partial \hat{\beta}_1^l} & \frac{\partial^2 L^*}{\partial \hat{\beta}_2^{l^2}} & \cdots & \frac{\partial^2 L^*}{\partial \hat{\beta}_2^l \partial \hat{\beta}_k^l} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L^*}{\partial \hat{\beta}_k^l \partial \hat{\beta}_1^l} & \frac{\partial^2 L^*}{\partial \hat{\beta}_k^l \partial \hat{\beta}_2^l} & \cdots & \frac{\partial^2 L^*}{\partial \hat{\beta}_k^{l^2}} \end{bmatrix} \quad (2.21)$$

である。(2.17) 式より $\nabla^2 L^*(\hat{\beta}^l)$ を求めると

$$\nabla L^*(\hat{\beta}^l) = J'(\hat{\beta}^l) \hat{\Sigma}^{-1} f(\hat{\beta}^l) \quad (2.22)$$

$$\begin{aligned} \nabla^2 L^*(\hat{\beta}^l) &= \nabla \{J'(\hat{\beta}^l) \hat{\Sigma}^{-1} f(\hat{\beta}^l)\} \\ &= J'(\hat{\beta}^l) \hat{\Sigma}^{-1} J(\hat{\beta}^l) + \sum_{i=1}^n f(\hat{\beta}^l) \hat{\Sigma}^{-1} \nabla^2 f(\hat{\beta}^l) \end{aligned} \quad (2.23)$$

となる。ここで $\hat{\beta}$ に高い精度を要求しても最終的に求める選択率 $\hat{p}_{it}$ の精度が高々 10^{-2} 程度であるので (2.23) 式の2次微分の右辺第2項を0に近いと見做して

$$\nabla^2 L^*(\hat{\beta}^l) = J'(\hat{\beta}^l) \hat{\Sigma}^{-1} J(\hat{\beta}^l) \quad (2.24)$$

とするガウス・ニュートン法を用いることにする。

(2.22) 式と (2.24) 式を点 $\hat{\beta}^l$ において考えると (2.19) 式は

$$\begin{aligned} \hat{\beta}^{l+1} &= \beta^l - \{\nabla^2 L^*(\hat{\beta}^l)\}^{-1} \nabla L^*(\hat{\beta}^l) \\ &= \beta^l - \{J'(\hat{\beta}^l) \hat{\Sigma}^{-1} J(\hat{\beta}^l)\}^{-1} J'(\hat{\beta}^l) \hat{\Sigma}^{-1} f(\hat{\beta}^l) \end{aligned} \quad (2.25)$$

となる。これを次の停止基準 (termination criterion) を満たすまで $l=1, 2, \dots$ について行う。

$$\frac{|\hat{\beta}^{l+1} - \hat{\beta}^l|}{|\hat{\beta}^l|} \leq \varepsilon \quad \varepsilon: \text{収束判定定数}$$

次にJacobian Matrix $J(\hat{\beta})$ を求める。式の煩雑さをさけるため、特に $m=2$ の場合について展開する。 $i=1$ に対しての $\nabla f(\hat{\beta})$ は

$$\frac{\partial \hat{f}_1(\hat{\beta})}{\partial \hat{\beta}_1} = \frac{\exp(\hat{\beta}' x_{1t}) \exp(\hat{\beta}' x_{2t}) (x_{2t1} - x_{1t1})}{\{\exp(\hat{\beta}' x_{1t}) + \exp(\hat{\beta}' x_{2t})\}^2} = (x_{2t1} - x_{1t1}) \hat{p}_{1t} \hat{p}_{2t} \quad (2.26)$$

と微分される。

同様に

$$\left. \begin{aligned} \frac{\partial \hat{f}_1(\hat{\beta})}{\partial \hat{\beta}_2} &= (x_{2t2} - x_{1t2}) \hat{p}_{1t} \hat{p}_{2t} \\ \frac{\partial \hat{f}_1(\hat{\beta})}{\partial \hat{\beta}_3} &= (x_{2t3} - x_{1t3}) \hat{p}_{1t} \hat{p}_{2t} \end{aligned} \right\} \quad (2.27)$$

と微分される。(2.26)、(2.27) 式から帰納的に次式を得る。

$$\frac{\partial \hat{f}_{1t}(\hat{\beta})}{\partial \hat{\beta}_j} = (x_{2tj} - x_{1tj}) \hat{p}_{1t} \hat{p}_{2t} \quad (j=1,2,\dots,k) \quad (2.28)$$

$i=2$ に対する微分は

$$\frac{\partial \hat{f}_2(\hat{\beta})}{\partial \hat{\beta}_1} = \frac{\exp(\hat{\beta}' x_{1t}) \exp(\hat{\beta}' x_{2t}) (x_{1t1} - x_{2t1})}{\{\exp(\hat{\beta}' x_{1t}) + \exp(\hat{\beta}' x_{2t})\}^2} = -(x_{2t1} - x_{1t1}) \hat{p}_{1t} \hat{p}_{2t} \quad (2.29)$$

となる。同様に

$$\left. \begin{aligned} \frac{\partial \hat{f}_2(\hat{\beta})}{\partial \hat{\beta}_2} &= -(x_{2t2} - x_{1t2}) \hat{p}_{1t} \hat{p}_{2t} \\ \frac{\partial \hat{f}_2(\hat{\beta})}{\partial \hat{\beta}_3} &= -(x_{2t3} - x_{1t3}) \hat{p}_{1t} \hat{p}_{2t} \end{aligned} \right\} \quad (2.30)$$

と微分される。(2.29)、(2.30) 式を一般化すると

$$\frac{\partial \hat{f}_2(\hat{\beta})}{\partial \hat{\beta}_j} = -(x_{2tj} - x_{1tj}) \hat{p}_{1t} \hat{p}_{2t} \quad (j=1,2,\dots,k) \quad (2.31)$$

となる。さらに、(2.28)、(2.31) 式をまとめると

$$\frac{\partial \hat{f}_i(\hat{\beta})}{\partial \hat{\beta}_j} = (-1)^{i+1} (x_{2tj} - x_{1tj}) \hat{p}_{1t}(\hat{\beta}) \hat{p}_{2t}(\hat{\beta}) \quad (j=1,2,\dots,k) \quad (2.32)$$

が得られる。これらの結果より $\nabla f_i(\hat{\beta})$ による $J_i(\hat{\beta})$ は

$$J_i(\hat{\beta}) = (-1)^{i+1} \begin{bmatrix} \hat{p}_{11} \hat{p}_{21} (x_{211} - x_{111}) & \hat{p}_{11} \hat{p}_{21} (x_{212} - x_{112}) & \cdots & \hat{p}_{11} \hat{p}_{21} (x_{21k} - x_{11k}) \\ \hat{p}_{12} \hat{p}_{22} (x_{221} - x_{121}) & \hat{p}_{12} \hat{p}_{22} (x_{222} - x_{122}) & \cdots & \hat{p}_{12} \hat{p}_{22} (x_{22k} - x_{12k}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{p}_{1n} \hat{p}_{2n} (x_{2n1} - x_{1n1}) & \hat{p}_{1n} \hat{p}_{2n} (x_{2n2} - x_{1n2}) & \cdots & \hat{p}_{1n} \hat{p}_{2n} (x_{2nk} - x_{1nk}) \end{bmatrix} \quad (i=1,2)$$

(2.33)

となり、 $J(\hat{\beta})$ は

$$J(\hat{\beta}) = \begin{bmatrix} J_1(\hat{\beta}) \\ J_2(\hat{\beta}) \end{bmatrix} \quad (2.34)$$

である。

 $J(\hat{\beta})$ を用いると $\nabla L^*(\hat{\beta})$ は

$$\nabla L^*(\hat{\beta}) = J'(\hat{\beta}) \hat{\Sigma}^{-1} \hat{f}(\hat{\beta}) = \begin{bmatrix} \sum_{i=1}^2 \sum_{t=1}^n (x_{1t1} - x_{2t1})(p_{it} - \hat{p}_{it}) \\ \sum_{i=1}^2 \sum_{t=1}^n (x_{1t2} - x_{2t2})(p_{it} - \hat{p}_{it}) \\ \vdots \\ \sum_{i=1}^2 \sum_{t=1}^n (x_{1tk} - x_{2tk})(p_{it} - \hat{p}_{it}) \end{bmatrix} \quad (2.35)$$

と表すことができる。同様に $\nabla^2 L^*(\hat{\beta})$ は

$$\nabla^2 L^*(\hat{\beta}) = J'(\hat{\beta}) \hat{\Sigma}^{-1} J(\hat{\beta}) = \begin{bmatrix} \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2t1} - x_{1t1})^2 & \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2t1} - x_{1t1})(x_{2t2} - x_{1t2}) & \cdots & \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2t1} - x_{1t1})(x_{2tk} - x_{1tk}) \\ \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2t2} - x_{1t2})(x_{2t1} - x_{1t1}) & \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2t2} - x_{1t2})^2 & & \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2t2} - x_{1t2})(x_{2tk} - x_{1tk}) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2tk} - x_{1tk})(x_{2t1} - x_{1t1}) & \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2tk} - x_{1tk})(x_{2t2} - x_{1t2}) & \cdots & \sum_{i=1}^2 \sum_{t=1}^n \hat{p}_{1t} \hat{p}_{2t} (x_{2tk} - x_{1tk})^2 \end{bmatrix} \quad (2.36)$$

ここで、(2.35)、(2.36) 式の $\nabla L^*, \nabla^2 L^*$ を用いて (2.25) 式を表すと

$$\hat{\beta}^{l+1} = \hat{\beta}^l - \{\nabla^2 L^*(\hat{\beta}^l)\}^{-1} \nabla L^*(\hat{\beta}^l) \quad (l = 0, 1, 2, \dots)$$

となる。これより数列 $\{\hat{\beta}^l\}$ を停止基準を満たすまで求めることができる。

3 関数 $\hat{p}_{it}(\hat{\beta})$ の離散化

離散型選択モデル

$$\hat{p}_{it} = \frac{\exp(\hat{\beta}'_i x_{it})}{\exp(\hat{\beta}'_1 x_{1t}) + \exp(\hat{\beta}'_2 x_{2t}) + \cdots + \exp(\hat{\beta}'_m x_{mt})} \quad \begin{matrix} (t=1,2,\cdots,n) \\ (i=1,2,\cdots,m) \end{matrix}$$

は定義域 $(-\infty, \infty)$ 、値域 $(0,1)$ の連続関数である。形状は図1に示すとおり、1/2で変曲点、1を漸近線にもつロジスティック曲線となる。

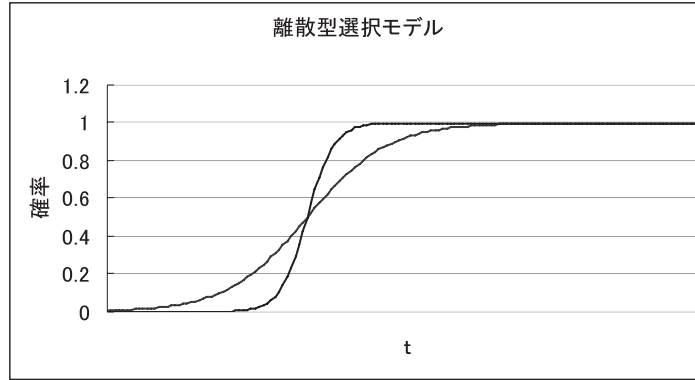


図1 離散型選択モデルの形状

このロジスティック曲線はパラメータ $\hat{\beta}$ の値により形状が変わる。

交通工学の分野では推定された $\hat{p}_{it}$ の値に従い次のように

$$\text{選択確率} = \begin{cases} 1 & \hat{p}_{it} \geq 0.5 \\ 0 & \hat{p}_{it} < 0.5 \end{cases}$$

離散化して選択確率を求めている。

4 パラメータ $\hat{\beta}$ の誤差が $\hat{p}_{it}$ に及ぼす影響

いま、 $\hat{p}_{it}$ と $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ の関係を

$$\hat{p}_{it} = \hat{f}_i(\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k) \quad (i=1,2,\cdots,m) \quad (4.1)$$

で表す。 $\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_k$ の絶対誤差を $\Delta\hat{\beta}_1, \Delta\hat{\beta}_2, \cdots, \Delta\hat{\beta}_k$ とすると、 $\hat{p}_{it}$ には $\Delta\hat{p}_{it}$ の誤差が生ずるのでこれらの関係を

$$\hat{p}_{it} + \Delta\hat{p}_{it} = \hat{f}_i(\hat{\beta}_1 + \Delta\hat{\beta}_1, \hat{\beta}_2 + \Delta\hat{\beta}_2, \cdots, \hat{\beta}_k + \Delta\hat{\beta}_k) \quad (i=1,2,\cdots,m) \quad (4.2)$$

で表し、 $\hat{p}_{it}$ のまわりにTaylor展開すると

$$\begin{aligned}\hat{p}_{it} + \Delta \hat{p}_{it} = \hat{f}_i(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k) + \Delta \hat{\beta}_1 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_1} + \Delta \hat{\beta}_2 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_2} + \dots + \Delta \hat{\beta}_k \frac{\partial \hat{f}_i}{\partial \hat{\beta}_k} \\ + \frac{1}{2!} \left(\Delta \hat{\beta}_1 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_1} + \Delta \hat{\beta}_2 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_2} + \dots + \Delta \hat{\beta}_k \frac{\partial \hat{f}_i}{\partial \hat{\beta}_k} \right)^2 \dots \quad (4.3)\end{aligned}$$

となるので、 $\Delta \hat{\beta}_1, \Delta \hat{\beta}_2, \dots, \Delta \hat{\beta}_k$ が非常に小さいと仮定し、2nd order以上を省略すると(4.3)式は近似的に

$$\hat{p}_{it} + \Delta \hat{p}_{it} = \hat{f}_i(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k) + \Delta \hat{\beta}_1 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_1} + \Delta \hat{\beta}_2 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_2} + \dots + \Delta \hat{\beta}_k \frac{\partial \hat{f}_i}{\partial \hat{\beta}_k} \quad (4.4)$$

と表すことができる。(4.1)、(4.2)、(4.4)式より絶対誤差 $\Delta \hat{p}_{it}$ は

$$\Delta \hat{p}_{it} = \Delta \hat{\beta}_1 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_1} + \Delta \hat{\beta}_2 \frac{\partial \hat{f}_i}{\partial \hat{\beta}_2} + \dots + \Delta \hat{\beta}_k \frac{\partial \hat{f}_i}{\partial \hat{\beta}_k} \quad (4.5)$$

と表すことができる。(4.1)式を $m=2$ の離散型選択モデル

$$\hat{p}_{it} = \frac{\exp(\hat{\beta}'_i x_{it})}{\exp(\hat{\beta}'_1 x_{1t}) + \exp(\hat{\beta}'_2 x_{2t})} \quad \begin{cases} (t=1,2,\dots,n) \\ (i=1,2) \end{cases} \quad (4.6)$$

で表すと(4.1)、(4.6)、(2.31)式より

$$\frac{\partial \hat{f}_i(\hat{\beta})}{\partial \hat{\beta}_j} = \frac{\partial \hat{p}_{it}}{\partial \hat{\beta}_j} = (-1)^j (x_{2jt} - x_{1jt}) \hat{p}_{1t} \hat{p}_{2t} \quad (j=1,2,\dots,k) \quad (4.7)$$

となるから、(4.7)式を各 $\hat{\beta}_i (i=1,2,\dots,k)$ について書き直すと

$$\left. \begin{aligned} \frac{\partial \hat{p}_{1t}}{\partial \hat{\beta}_1} &= -(x_{21t} - x_{11t}) \hat{p}_{1t} \hat{p}_{2t} \\ \frac{\partial \hat{p}_{1t}}{\partial \hat{\beta}_2} &= -(x_{22t} - x_{12t}) \hat{p}_{1t} \hat{p}_{2t} \\ &\dots\dots\dots \\ \frac{\partial \hat{p}_{1t}}{\partial \hat{\beta}_k} &= -(x_{2kt} - x_{1kt}) \hat{p}_{1t} \hat{p}_{2t} \end{aligned} \right\} \quad (4.8)$$

$$\left. \begin{aligned} \frac{\partial \hat{p}_{2t}}{\partial \hat{\beta}_1} &= (x_{2t1} - x_{1t1}) \hat{p}_{1t} \hat{p}_{2t} \\ \frac{\partial \hat{p}_{2t}}{\partial \hat{\beta}_2} &= (x_{2t2} - x_{1t2}) \hat{p}_{1t} \hat{p}_{2t} \\ &\dots\dots\dots \\ \frac{\partial \hat{p}_{2t}}{\partial \hat{\beta}_k} &= (x_{2tk} - x_{1tk}) \hat{p}_{1t} \hat{p}_{2t} \end{aligned} \right\} \quad (4.9)$$

が得られる。(4.5)、(4.6)、(4.8)、(4.9) 式より絶対誤差は

$$\left\{ \begin{aligned} \Delta \hat{p}_{1t} &= \Delta \hat{\beta}_1 (x_{1t1} - x_{2t1}) \hat{p}_{1t} \hat{p}_{2t} + \Delta \hat{\beta}_2 (x_{1t2} - x_{2t2}) \hat{p}_{1t} \hat{p}_{2t} + \dots + \Delta \hat{\beta}_k (x_{1tk} - x_{2tk}) \hat{p}_{1t} \hat{p}_{2t} \\ \Delta \hat{p}_{2t} &= \Delta \hat{\beta}_1 (x_{2t1} - x_{1t1}) \hat{p}_{1t} \hat{p}_{2t} + \Delta \hat{\beta}_2 (x_{2t2} - x_{1t2}) \hat{p}_{1t} \hat{p}_{2t} + \dots + \Delta \hat{\beta}_k (x_{2tk} - x_{1tk}) \hat{p}_{1t} \hat{p}_{2t} \end{aligned} \right. \quad (4.10)$$

となる。一般に $\Delta \hat{p}_{1t}, \Delta \hat{p}_{2t}$ は絶対値評価されている場合が多いので相対誤差 $\Delta \hat{p}_{1t} / \hat{p}_{1t}$ は

$$\begin{aligned} \left| \frac{\Delta \hat{p}_{1t}}{\hat{p}_{1t}} \right| &= \left| \frac{\Delta \hat{\beta}_1}{\hat{\beta}_1} \right| \frac{\partial \hat{p}_{1t}}{\partial \hat{\beta}_1} + \left| \frac{\Delta \hat{\beta}_2}{\hat{\beta}_2} \right| \frac{\partial \hat{p}_{1t}}{\partial \hat{\beta}_2} + \dots + \left| \frac{\Delta \hat{\beta}_k}{\hat{\beta}_k} \right| \frac{\partial \hat{p}_{1t}}{\partial \hat{\beta}_k} \\ &= \left| \frac{\Delta \hat{\beta}_1}{\hat{\beta}_1} \right| (x_{1t1} - x_{2t1}) + \left| \frac{\Delta \hat{\beta}_2}{\hat{\beta}_2} \right| (x_{1t2} - x_{2t2}) + \dots + \left| \frac{\Delta \hat{\beta}_k}{\hat{\beta}_k} \right| (x_{1tk} - x_{2tk}) \end{aligned} \quad (4.11)$$

$$\begin{aligned} \left| \frac{\Delta \hat{p}_{2t}}{\hat{p}_{2t}} \right| &= \left| \frac{\Delta \hat{\beta}_1}{\hat{\beta}_1} \right| \frac{\partial \hat{p}_{2t}}{\partial \hat{\beta}_1} + \left| \frac{\Delta \hat{\beta}_2}{\hat{\beta}_2} \right| \frac{\partial \hat{p}_{2t}}{\partial \hat{\beta}_2} + \dots + \left| \frac{\Delta \hat{\beta}_k}{\hat{\beta}_k} \right| \frac{\partial \hat{p}_{2t}}{\partial \hat{\beta}_k} \\ &= \left| \frac{\Delta \hat{\beta}_1}{\hat{\beta}_1} \right| (x_{2t1} - x_{1t1}) + \left| \frac{\Delta \hat{\beta}_2}{\hat{\beta}_2} \right| (x_{2t2} - x_{1t2}) + \dots + \left| \frac{\Delta \hat{\beta}_k}{\hat{\beta}_k} \right| (x_{2tk} - x_{1tk}) \end{aligned} \quad (4.12)$$

として表すのが一般的である。全体のおおよその相対誤差を把握するため

$$\text{standard error} = \sqrt{\frac{1}{2n} \sum_{t=1}^2 \sum_{i=1}^n \left(\frac{\Delta \hat{p}_{it}}{\hat{p}_{it}} \right)^2} \quad (4.13)$$

を用いる。

5 モデル適用例

前節までの方法を北海道本州間の貨物輸送における輸送手段選択モデルのパラメータ推定に適用した。用いたデータは平成12年1月に北海道内の荷主 ($n = 23$) を対象に、利用輸送手段（鉄道、フェリー）、運賃、所要時間について実態調査を行ったものである。

輸送機関選択モデルは

$$\hat{p}_{it} = \frac{\exp(\hat{\beta}' x_{it})}{\exp(\hat{\beta}' x_{1t}) + \exp(\hat{\beta}' x_{2t})} \quad \left\{ \begin{array}{l} (t = 1, 2, \dots, 23) \\ i = 1: \text{鉄道} \\ i = 2: \text{フェリー} \end{array} \right.$$

$$\hat{\beta}' x_{it} = \hat{\beta}_1 x_{it1} + \hat{\beta}_2 x_{it2}$$

$\hat{p}_{it}$ ：選択確率、 $\hat{\beta}_1$ ：運賃に対するパラメータ、 $\hat{\beta}_2$ ：所要時間に対するパラメータである。最尤法によるパラメータ推定結果と選択確率の推定結果を表1、表2、表3に示す。

表1 最尤法によるパラメータ推定結果

反復数 l	パラメータ $\hat{\beta}_1$	相対誤差 $ \hat{\beta}_1^{l+1} - \hat{\beta}_1^l / \hat{\beta}_1^l $	パラメータ $\hat{\beta}_2$	相対誤差 $ \hat{\beta}_2^{l+1} - \hat{\beta}_2^l / \hat{\beta}_2^l $
初期値	0		0	
1	-2.194144	2.194145	-0.359133	0.359133
2	-3.635860	0.657074	-0.592087	0.648658
3	-4.880005	0.342187	-0.781481	0.319876
4	-5.586803	0.144835	-2.882770	0.129611
5	-5.745291	0.028368	-0.904460	0.024570
6	-5.751507	0.001082	-0.905279	0.000906
7	-5.751516	0.000002	-0.905280	0.000001
8	-5.751516	0.000000	-0.905280	0.000000

表2 鉄道選択確率の推定結果

荷主番号	観測値	推定値	
1	0	0	
2	0	0	
3	1	1	
4	0	0	
5	0	0	
6	0	0	
7	1	1	
8	0	0	
9	0	0	
10	0	0	
11	0	0	
12	0	0	
13	1	1	
14	1	1	
15	0	1	*
16	0	0	
17	0	0	
18	1	1	
19	0	0	
20	0	1	*
21	1	1	
22	0	0	
23	0	0	

表3 フェリー選択確率の推定結果

荷主番号	観測値	推定値	
1	1	1	
2	1	1	
3	0	0	
4	1	1	
5	1	1	
6	1	1	
7	0	0	
8	1	1	
9	1	1	
10	1	1	
11	1	1	
12	1	1	
13	0	0	
14	0	0	
15	1	0	*
16	1	1	
17	1	1	
18	0	0	
19	1	1	
20	1	0	*
21	0	0	
22	1	1	
23	1	1	

表2、表3における*は実測値と推定値が等しくない観測点を示す。

$\hat{\beta}$ が表1に示される誤差をもつ場合、4で述べた方法で $\hat{p}_i$ への伝播誤差を求めたところ選択率の伝播誤差= 0.4×10^{-12} であった。3で述べたとおり $\hat{p}_i$ の精度は高々 10^{-2} 程度であるので $\hat{\beta}$ の精度は十分であると思う。

6 おわりに

これまで離散型選択モデルは交通計画の多くの例に適用され、重要な役目を果たしてきている。しかし、その理論的背景にはいくつかの問題点があり、仮定の検証も行われずにモデルが適用されてきた。

本研究で示したエントロピー理論から導出した離散型選択モデルは、既存のモデルと同じ結果を与え、かつモデルモデル導出の理論的整合性が明確になったと考えている。

参考文献

- 1) 土木学会：非集計行動モデルの理論と実際、p16、1995
- 2) 林知己夫：数量化の方法、東洋経済新報社、1974
- 3) 交通工学研究会編：やさしい非集計分析、p25、1993
- 4) ビリングスレイ：渡辺毅・十時東生共訳：確率論とエントロピー、数学叢書4、

吉岡書店、1968

5) 今野浩・山下浩著：非線形計画法、日科技連、1978

6) 中川徹・小柳義夫著：非線形最小二乗法による実験データ解析、東京大学出版会、1982