ロジット・ワイビットモデルの一般化の改良

中山　晶一朗[1]・力石　真[2]

[1]正会員　金沢大学　環境デザイン学系（〒920-1192　金沢市角間町）
E-mail: snakayama@t.kanazawa-u.ac.jp

[2]正会員　広島大学大学院　国際協力研究科（〒739-8529　東広島市鏡山一丁目５番１号）
E-mail: chikaraishim@hiroshima-u.ac.jp

交通現象は人間行動である交通行動の集積によって構成されている．交通行動としては，複数の経路から実際の走行する経路を一つ選ぶなど交通行動には複数の選択肢から一つを選ぶという離散選択が多い．交通分野では，ランダム効用理論に基づいた離散選択モデルの中で，操作性の高さからロジットモデルが多用されている．近年，ワイブル誤差項をもつワイビットモデルも登場している．本稿では，ロジットモデルとワイビットモデルを統一的に一般化したモデルを提案する．

Key Words :*generalized logit model, Tsallis statistics*

1. はじめに

交通現象としては，個々の車両の挙動，それをマクロ的にとらえた交通流のみならず，個々の車両が利用されるのかどうか，利用されるとすると，どこからどこへ何時にどの経路で走行するのか，なども重要な問題である．交通流を考える場合に重要な要因である車両数は，人間行動である交通行動の集積によって構成されている．

朝の通勤時間などピーク時間では，多くの自動車利用者は自宅から職場へ向かう．利用者の出発時刻は起床時間等で変更がなされないと考え得る場合，交通行動としては，経路選択が主要な問題となろう．このように，複数の経路から実際の走行する経路を一つ選ぶという経路選択など交通行動には複数の選択肢から一つを選ぶという離散選択が多い．交通分野では，ランダム効用理論に基づいた離散選択モデルの中で，操作性の高さからロジットモデルが多用されている[1][2].

ロジットモデルは次章で詳述するが，ランダム効用理論に基づいた離散選択モデルの一つであり，効用が極値分布の一つであるガンベル分布に従う場合その離散選択モデルはロジットモデルと呼ばれる．効用が正規分布に従う場合はプロビットモデルと呼ばれている．中心極限定理などから類推できるように，プロビットモデルの方が現象の再現上本来的には望ましいとも考えられるが，正規分布の確率密度関数は積分ができないことを起因として，プロビットモデルは数値的に扱うもしくは近似的にしか扱うことができないという欠点を持つ．一方，次章で述べるようにロジットモデルは閉形式で問題を記述することができ，数理的な操作性が高く，交通分野ではプロビットモデルよりもロジットモデルが多用されている．

ロジットモデルで用いられているガンベル分布は極値分布の一つであるが，ガンベル分布を含むより一般的な分布として一般化極値分布がある．本研究では，ロジットモデルで用いられるガンベル分布を一般化極値分布に拡張した場合の離散選択モデルを導出する．そして，この導出されたモデルとツァリス統計との関連性について考察する．

2. ロジットモデル

ロジットモデルは，効用理論（ランダム効用理論）に基づいた個人の行動（選択）のモデル化が行われたものである．ランダム効用理論は，効用理論と同じく，その選択肢の望ましさを意味する効用が最も大きい選択肢を選ぶというものである．通常の効用理論との違いは，効用は確定的ではなく，確率的である点である．

本研究では，複数の選択肢から一つの選択肢を選ぶ離散選択問題を取り扱う．例えば，駅から大学まで行く場合の交通手段を決める問題としては，バス・地下鉄・タクシーの3つの選択肢から一つ選ぶという離散選択問題などがある．

上述の通り，効用は各選択肢の望ましさを意味する．

選択する個人が合理的であることを仮定すると，効用が最も大きい選択肢を選択する．これが効用理論である．ランダム効用理論はこの効用が確率的であるため，ある選択肢の効用が最も高くなることも確率的となる．ランダム効用理論に基づいた離散選択問題では，各選択肢について，その選択肢の効用が最も高くなる確率を求めることになる．これが各選択肢が選択される確率となる．

現在，交通の分野で用いられているランダム効用理論では，通常，効用は以下の式によって表される．

$$U_i = v_i + \varepsilon_i \tag{1}$$

ここで，U_i は選択肢 i の効用，v_i は選択肢 i の確定効用（効用の確定的な部分），ε_i は確率項（効用の確率部分）を表す．確率項の解釈については様々なものが考えられるが，本研究では，確定項の値を正確に知ることができず，選択肢の効用を正確に把握できないために付加されたものとみなす．

ある個人が選択肢1を選択するのは，選択肢1の効用 U_1 が他の選択肢の効用 $U_2, U_3, \ldots$ よりも大きい場合（$U_1 > U_2, U_1 > U_3, \ldots$）であり，選択肢1を選択する確率は

$$p_1 = \Pr[U_1 > U_2]\Pr[U_1 > U_3]\cdots\Pr[U_1 > U_I] \tag{2}$$

である．そして，これは，

$$p_1 = \Pr[U_1 > \max_{i(\neq 1)}(U_i)] \tag{3}$$

である．ここで，$\Pr[*]$ は * の生起する確率で，I は選択肢の総数，$\max(\cdot)$は最大値をとる演算子である．

上述のように誤差項は独立であることを考えると，

$$p_1 = \int_{x\in\Omega} \Pr[U_1 = x]\Pr[U_2 < x]\cdots\Pr[U_I < x]\,dx \tag{4}$$

そして，ε_i が従う確率分布の確率密度関数を $f_i(x_i)$，累積分布関数を $F_i(x_i)$ とすると，選択肢1が選択される確率は

$$p_1 = \int_{x\in\Omega} f_1(x)F_2(x)\cdots F_I(x)\,dx \tag{5}$$

となる．

当然のことながら，選択肢1が選択される確率 p_1 は $f_i(x_i)$ もしくは $F_i(x_i)$ が従う確率分布形によって異なる．交通の分野で誤差項の確率分布として最もよく用いられているのがガンベル分布である．誤差項が独立なガンベル分布に従う場合のランダム効用理論に基づく離散選択モデルはロジットモデルと呼ばれている．誤差項が標準ガンベル分布に従うとする．

標準ガンベル分布の累積分布関数は

$$F(x) = \exp[-\exp(-x)] \tag{6}$$

である．ガンベル分布は二重指数分布とも呼ばれ，極値分布の一つである．極値分布I型やガンベル型極値分布とも呼ばれている．

誤差項 ε_i が標準ガンベル分布に従う場合，ランダム効用 $U_i = v_i + \varepsilon_i$ の累積分布関数は定式を用いて

$$\begin{aligned} F_i(x) &= \exp[-\exp\{-(x - v_i)\}] \\ &= \exp[-\exp(v_i)\exp(-x)] \end{aligned} \tag{7}$$

そして，その確率密度関数は

$$\begin{aligned} f_i(x) &= \frac{dF_i(x)}{dx} = \frac{d}{dx}\exp[-\exp\{-(x - v_i)\}] \\ &= \exp[-\exp\{-(x - v_i)\}]\frac{d}{dx}[-\exp\{-(x - v_i)\}] \\ &= \exp[-(x - v_i)]F_i(x) \\ &= \exp(v_i)\exp(-x)F_i(x) \end{aligned} \tag{8}$$

これを式(5)に入れると，

$$\begin{aligned} p_1 &= \exp(v_1)\int_{-\infty}^{\infty}\exp(-x)F_1(x)F_2(x)\cdots F_I(x)\,dx \\ &= \exp(v_1)\int_{-\infty}^{\infty}\exp(-x)\exp\left[-\left(\sum_{i=1}^{I}\exp(v_i)\right)\exp(-x)\right]dx \end{aligned} \tag{9}$$

ここで，$z = \exp(-x)$ として，置換積分を行うと，

$$\begin{aligned} p_1 &= \exp(v_1)\int_0^{\infty}\exp\left[-\left(\sum_{i=1}^{I}\exp(v_i)\right)z\right]dz \\ &= \frac{\exp(v_1)}{\sum_{i=1}^{I}\exp(v_i)}\left[-\exp\left\{-\left(\sum_{i=1}^{I}\exp(v_i)\right)z\right\}\right]_{z=0}^{\infty} \\ &= \frac{\exp(v_1)}{\sum_{i=1}^{I}\exp(v_i)} \end{aligned} \tag{10}$$

これは任意の選択肢 i について成り立つため，選択肢 i が選択される確率は以下の通りとなる．

$$p_i = \frac{\exp(v_i)}{\sum_{i'=1}^{I}\exp(v_{i'})} \tag{11}$$

3. q–指数関数で表示された一般化極値分布

上述の通り，ロジットモデルでは，効用（もしくは誤差項）はガンベル分布に従っている．ガンベル分布は極値分布の1つのタイプであり，ガンベル分布を含むより広い極値分布には一般化極値分布がある．本章では，効用が一般化極値分布に従う場合のランダム効用理論に基づいた離散選択モデルを導出する．一般化極値分布はガンベル分布を一般化したものであり，このモデルはロジットモデルの一般化と言える．そして，このロジットモデルの一般化はツァリス統計と深い関係性を次章以降で明らかにする．

一般化極値分布の累積分布関数 $G(x)$ は以下の通りである．

$$\tilde{G}(x) = \exp\left\{-\left[1+\gamma\left(\frac{x-\mu}{\rho}\right)\right]^{-\frac{1}{\gamma}}\right\} \tag{12}$$

ここで，$\gamma = 0$ の時，$\exp(x) = \mathrm{Lim}_{\gamma\to 0}[1+\gamma x]^{\frac{1}{\gamma}}$ となるため，上式は

$$\tilde{G}(x) = \exp\left\{-\exp\left[-\left(\frac{x-\mu}{\rho}\right)\right]\right\} \tag{13}$$

となり，ガンベル分布となる．なお，さらに，$\rho = 1, \mu = 0$ の時は標準ガンベル分布となる．

本稿では，ツァリス統計力学との関連が分かるように，以下のq-指数関数

$$\exp_q(x) := [1+(1-q)x]^{\frac{1}{1-q}} \tag{14}$$

を用いて，一般化極値分布の累積分布関数 $G(x)$ を以下のように表現する．

$$\begin{aligned} G(x) &= \exp\left[-\exp_q\left(-\frac{x-\mu}{\rho}\right)\right] \\ &= \exp\left\{-\left[1+(1-q)\left(-\frac{x-\mu}{\rho}\right)\right]^{\frac{1}{1-q}}\right\} \end{aligned} \tag{15}$$

選択肢 i の確定効用を v_i とおく．そして，確定効用の関数 $h(v_i)$ を考える．後述の離散選択モデルの導出の準備のために，

$$G_i(x) = \exp[-h(v_i)\exp_q(-x)] \tag{16}$$

を導入する．

$$\begin{aligned} G_i(x) &= \exp[-h(v_i)\exp_q(-x)] \\ &= \exp\left[-\exp_q\left(-\frac{x-\frac{1-h^{q-1}(v_i)}{1-q}}{h^{q-1}(v_i)}\right)\right] \end{aligned} \tag{16}$$

であるため，この導入したものは一般化極値分布（の特殊形）である．なお，$h^{q-1}(v_i) = \{h(v_i)\}^{q-1}$ とする．

ここで，累積分布関数は $0 \le G_i(x) \le 1$ であるため，x の定義域は

$$\begin{cases} x \ge \dfrac{1}{1-q} & \text{if } q > 1 \\ x \le \dfrac{1}{1-q} & \text{if } q < 1 \end{cases} \tag{17}$$

となる．なお，$q = 1$ の時の定義域はガンベル分布と同様に実数全体である．

4. ロジットモデルのツァリス的一般化

ロジットモデルの場合と同様に各選択肢の効用は互いに独立であると仮定する．そして，累積分布関数 $G_i(x)$ とそれを微分した確率密度関数 $g_i(x) = dG_i(x)/dx$ を式(5)に代入すると，選択肢1が選択される確率は

$$p_1 = \int_{x\in\Omega} g_1(x)G_2(x)\cdots G_I(x)\,dx \tag{18}$$

として与えられる．そして，

$$g_1(x) = \frac{d}{dx}G_1(x) = G_1(x)h(v_1)[\exp_q(-x)]^q \tag{19}$$

であるため，

$$\begin{aligned} p_1 &= h(v_i)\int_{x\in\Omega}[\exp_q(-x)]^q G_1(x)G_2(x)\cdots G_I(x)\,dx \\ &= h(v_i)\int_{x\in\Omega}[\exp_q(-x)]^q \exp\left[-\left\{\sum_{i=1}^{I}h(v_i)\right\}\exp_q(-x)\right]dx \end{aligned} \tag{20}$$

ここで，

$$z := \exp_q(-x) \tag{21}$$

として，置換積分を行うと，

$$\begin{aligned} p_1 &= -h(v_1)\int_0^\infty \exp\left(z\sum_{i=1}^{I}h(v_i)\right)dz \\ &= \frac{h(v_1)}{\sum_{i=1}^{I}h(v_i)} \end{aligned} \tag{22}$$

これは任意の選択肢について成立するため，選択肢 i (=1, 2,…, I) が選択される確率 p_i は

$$p_i = \frac{h(v_1)}{\sum_{i=1}^{I}h(v_i)} \tag{23}$$

として与えられる．

選択肢 i の確率効用 U_i の平均は確定効用である v_i とする．このように平均が v_i となるためには，$G_i(x) = \exp[-h(v_i)\exp_q(-x)]$ の $h(v_i)$ は

$$h(v_i) = \frac{\exp_{2-q}(v_1)}{\exp_{2-q}[\Gamma(2-q)]} \tag{24}$$

と設定しなければならない．これを式23に入れると，

$$p_i = \frac{\exp_{2-q}(v_1)}{\sum_{i=1}^{I}\exp_{2-q}(v_i)} \tag{25}$$

が導かれる．

5. ワイビットモデルとの関連

誤差項がワイブル分布に従う離散選択モデルであるワイビットモデル[4)]は以下の通りである．

$$p_i = \frac{v_i^{\ \lambda}}{\sum_{i'=1}^{I} v_{i'}^{\ \lambda}} \tag{26}$$

q指数関数の性質から，定数 κ について

$$\exp_q(x+\kappa) = \exp_q(\kappa)\ \exp_q\left(\frac{x}{1+(1-q)\kappa}\right) \tag{27}$$

である．よって，

$$\exp_{2-q}(-1)\ \exp_{2-q}\left(\frac{v_i}{2-q}\right) = \exp_{2-q}\left(v_i - 1\right) = \left[(q-1)\hat{v}_i\right]^{\frac{1}{q-1}} \tag{28}$$

となる．これを式(27)に代入すると，

$$p_i = \frac{v_i^{\ 1/(q-1)}}{\sum_{i'=1}^{I} v_{i'}^{\ 1/(q-1)}} \tag{29}$$

が得られる．これは式(30)のワイビットモデルと同じである．よって，数値的には，本研究のモデルはワイビットモデルも含んでいることが分かる．

参考文献

1) 北村隆一，森川高行：交通行動の分析とモデリング，技報堂出版，東京，2002.
2) Train, K.E.: Discrete Choice Models with Simulation, Cambridge University Press, Cambridge, U.K., 2003
3) 須鎗弘樹：複雑系のための基礎数理：べき乗則とツァリスエントロピーの数理，牧野書店，東京，2010.
4) Castillo, E., Menendez, J. M., Jimenez, P., & Rivas, A.: Closed form expressions for choice probabilities in the Weibull case. Transportation Research, 42B, 373–80, 2008.

（2014. 8. 1 受付）