ロジット型交通選択モデルとツァリス統計

中山　晶一朗[1]

[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

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．本稿では，ロジットモデルをツァリス統計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．

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

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

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

となり，ガンベル分布となる．なお，さらに，$\theta=1, \mu=0$ 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

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

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．

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

なお，式(12)からは$\gamma=q-1, \mu=v, \theta=s-(1-q)v$ とパラメータ変換を施している．

各選択肢の効用は一般化極値分布に従うとする．選択肢 i の効用 U_i が従う一般化極値分布 $G_i(x)$ は以下の通りとする．

$$G_i(x)=\exp\left[-\exp_q\left(-\frac{x-v_i}{s-(1-q)v_i}\right)\right] \tag{16}$$

$q=1$の時のq-指数関数は通常の指数関数となる，つまり，$\exp_1(x)=\exp(x)$ のため，$q=1$, $s=1$ の時，$G_i(x)=\exp[-\exp\{-(x-v_i)\}]$ となり，前章で述べたロジットモデルとなる．

ここで，

$$\begin{aligned} &\exp_q\left(-\frac{x-v_i}{s-(1-q)v_i}\right)= \\ &\left[1+(1-q)\left(\frac{v_i}{s-(1-q)v_i}\right)+(1-q)\left(-\frac{x}{s-(1-q)v_i}\right)\right]^{\frac{1}{1-q}} \\ &=\left[1+(1-q)\left(\frac{v_i}{s-(1-q)v_i}\right)\right]^{\frac{1}{1-q}}\left[1+(1-q)\left(-\frac{x}{s}\right)\right]^{\frac{1}{1-q}} \\ &=\exp_q\left(\frac{v_i}{s-(1-q)v_i}\right)\exp_q\left(-\frac{x}{s}\right) \end{aligned} \tag{17}$$

である．そして，

$$\hat{v}_i:=\frac{v_i}{s-(1-q)v_i} \tag{18}$$

とすると，式(16)の $G_i(x)$ は以下のように書きなおすことができる．

$$G_i(x)=\exp\left[-\exp_q\left(\hat{v}_i\right)\exp_q\left(-\frac{x}{s}\right)\right] \tag{19}$$

$\exp_q\left(\hat{v}_i\right)$は$v_i/[s-(1-q)v_i]$のべき乗であるため，後の利便性を考慮し，本研究では，$\exp_q\left(\hat{v}_i\right)>0$ の場合のみを対象とする．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

$$\begin{cases} x\ge \dfrac{s}{1-q} & \text{if } q>1 \\ x\le \dfrac{s}{1-q} & \text{if } q<1 \end{cases} \tag{20}$$

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

$$p_1=\int_{x\in\Omega} g_1(x)G_2(x)\cdots G_I(x)\,dx \tag{21}$$

として与えられる．そして，

$$g_1(x)=\frac{d}{dx}G_1(x)=-\exp_q(\hat{v}_1)G_1(x)\frac{d}{dx}\left[\exp_q\left(-\frac{x}{s}\right)\right] \tag{22}$$

であり，ここで，後の置換積分の準備という意味も含め，

$$z:=\exp_q\left(-\frac{x}{s}\right) \tag{23}$$

とおくと，

$$g_1(x)=\frac{d}{dx}G_1(x)=-\exp_q(\hat{v}_1)G_1(x)\frac{dz}{dx} \tag{24}$$

となる．これを式(21)に代入すると，

$$p_1=-\exp_q(\hat{v}_1)\int_{z\in\hat{\Omega}} G_1(x)G_2(x)\cdots G_I(x)\,dz \tag{25}$$

となる．ここで，

$$\sigma=\sum_{i=1}^{I}\exp_q(\hat{v}_i) \tag{26}$$

とすると，

$$\prod_{i=1}^{I}G_i(x)=\exp\left[-\sigma\exp_q\left(-\frac{x}{s}\right)\right]=\exp(-\sigma z) \tag{27}$$

である．これを用いると，

$$\begin{aligned} p_1&=-\exp_q(\hat{v}_1)\int_{z\in\hat{\Omega}}\exp(-\sigma z)\,dz \\ &=\frac{\exp_q(\hat{v}_1)}{\sigma}\left[\exp(-\sigma z)\right]_{\infty}^{0} \\ &=\frac{\exp_q(\hat{v}_1)}{\sum_{i=1}^{I}\exp_q(\hat{v}_i)} \end{aligned} \tag{28}$$

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

$$p_i = \frac{\exp_q(\hat{v}_i)}{\sum_{i'=1}^{I} \exp_q(\hat{v}_{i'})} \tag{29}$$

として与えられる．

q = 1 の時は，明らかに式(29)は通常のロジットモデル式の式(11)であることが分かり，ロジットモデルの一般化であることを確認できる．

5. 一般化ロジットモデルとツァリスエントロピー

本章では，前章で導出された一般化極値分布に従った l効用に基づく離散選択モデル，つまり，一般化ロジットモデルとツァリス統計との関連性について考察する．ツァリスエントロピーは以下の通りとされている．

$$S_q(\mathbf{p}) = -\frac{1 - \sum_{i=1}^{I} p_i{}^q}{1-q} = -\sum_{i=1}^{I} p_i{}^q \ln_q(p_i) \tag{30}$$

ここで，$\mathbf{p}$ は p_i を要素に持つベクトル，$\ln_q(x)$ はq-対数関数で，

$$\ln_q(x) := \frac{x^{1-q} - 1}{1-q} \tag{31}$$

である．なお，$\ln_q(x)$ は，q = 1 の時，通常の自然対数関数，すなわち，$\ln_1(x) = \ln(x)$ である．また，$\ln_q(\exp_q[x]) = x$ である．

以下のような制約条件付きツァリスエントロピー最大化問題を考える．

$$\max_{\mathbf{p}}.\ S_q(\mathbf{p}) \tag{32}$$

$$s.t.\ \sum_{i=1}^{I} p_i = 1 \tag{33}$$

上記の最大化問題の解は，以下のラグランジュ関数

$$L_q(\mathbf{p}, \alpha, \beta) = S_q(\mathbf{p}) - \alpha\left(\sum_{i=1}^{I} p_i - 1\right) \tag{34}$$

によって，以下のように与えられる．

$$p_i^* = \frac{\exp_q(\hat{v}_i)}{\sum_{i'=1}^{I} \exp_q(\hat{v}_{i'})} \tag{35}$$

以上のように，ツァリスエントロピーを式(33)の下で最大化させせることによって，一般化極値分布に従って確率変動する効用に基づいた離散選択モデル（一般化ロジットモデル）の式(29)を導出することができ，ツァリスエントロピーと一般化ロジットモデルの関連が分かる．このように，ガンベル分布を一般化させた一般化極値分布を用いるというロジットモデルの一般化はツァリス統計と関連が深いことが明らかになり，本研究では，このような一般化をロジットモデルのツァリス的一般化と呼ぶことにする．

6. 配分モデル

通常のロジット型の確率的利用者均衡は，通常のエントロピー項を含んだ最小化問題として定式化することができる．これを通常のエントロピーをツァリスエントロピーに置き換えた以下の最小化問題を考えよう．

$$\min_{\mathbf{p}}.\ \sum_{i=1}^{I}\sum_{j=1}^{J_i} p_{ij}{}^q c_{ij} - \frac{1}{\theta}\sum_{i=1}^{I} S_q(\mathbf{p}_i) \tag{32}$$

$$s.t.\ \sum_{i=1}^{I} p_i = 1 \tag{33}$$

$$L_q(\mathbf{p}, \boldsymbol{\lambda}) = \sum_{i=1}^{I}\sum_{j=1}^{J_i} p_{ij}{}^q c_{ij} - \frac{1}{\theta}\sum_{i=1}^{I} S_q(\mathbf{p}_i) - \sum_{i=1}^{I} \lambda_i \left(\sum_{j=1}^{J_i} p_{ij} - 1\right)$$

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.

（2012. 8. 3 受付）

LOGIT-TYPE TRAVEL CHOICE MODELS AND TSALLIS STATISTICS

Sho-ichiro NAKAYAMA

Traffic or transportation phenomena consist of travel behaviors as human behaviors. There are many

discrete choice problems in which a single one is chosen of several alternatives such as route choice. The logit model is widely used in the field of travel behavior analysis because of its easy mathematical handling. In this study, the logit model is generalized from the standpoint of Tsallis statistics.