

Daniel McFadden (1978)

Modeling the Choice of Residential Location

理論談話会 #2 行動モデル I

5月19日

B4 月田 光

1. 目的
2. 住宅立地選択論
3. MNLモデル
4. NLモデル
5. GEVモデル
6. NLモデルとGEVモデル
7. 選択肢数の制限
8. 選択肢の集約と「類似性」の効果

- 経済的選択行動の理論を、住宅立地の実証分析に適した具体的なモデルに変換する問題。特に、住宅立地間の個人の選択の問題

→その中でも、モデル化における二つの問題を扱う

1. 一般的なMNLモデルで想定していない「類似性」への対応
~住居が類似しているという認識が選択に与える影響を考える。

参考1) IIA特性(independence from irrelevant alternatives条件)

→選択肢同士が互いに独立だというMNLモデルにおける条件

参考2) 青バス赤バス問題 →類似性を説明する際によく使われる例え

本論文では確率的選択モデルの一群(family)を紹介する。

- 本モデルでは類似すると知覚される住戸は集約される。
- 従来のMNLモデルはその中の特殊なケースとなる。

→パラメータの振り方により、一方の極端な値にするとMNLモデルに、もう一方の極値にすると同じ属性の住戸は同じものとして認識。

- 経済的選択行動の理論を、住宅立地の実証分析に適した具体的なモデルに変換する問題。特に、住宅立地間の個人の選択の問題

→その中でも、モデル化における二つの問題を扱う

2. 選択肢の数が非現実的に多い場合の選択モデルの推定

~選択肢の数を制限する場面では、推定メカニズムを適切に調整して、選択肢の全集合からサンプリングすることによってモデルを推定する。

→これにより、限られたデータ収集と計算で推定ができる。

前提：消費者は合理的で効用を最大化する(古典的なモデル)

コミュニティ $c=1,2,\dots,C$
c内の住戸 $n=1,2,\dots,N_c$

選択肢cnが持つ効用 U_{cn}

内容：選択肢の属性(アクセス、公共サービスの質、近隣など)
消費者の特性(年齢、家族の大きさ、所得など)

効用 U_{cn} を次の二つに分解

$$U_{cn} = V_{cn} + \varepsilon_{cn}$$

V_{cn} ：観測可能な変数のベクトルで指定された関数である項

ε_{cn} ：観測できない変数の寄与度をまとめた項

選択肢全ての属性が観測されるわけではなく、観測できない変数もある。
そしてこれは観測された変数で条件付けられた何らかの確率分布を持つ。

もし観察者が効用関数の形と観察されていない変数の確率分布を知っていれば、選択肢の期待分布について確率論的な記述を行うことができる

$$P_{cn} = \text{Prob}[U_{cn} > U_{bm} \text{ for } bm \neq cn] \quad (1)$$

$\vec{\varepsilon} = (\varepsilon_{11}, \dots, \varepsilon_{1N_1}, \dots, \varepsilon_{c1}, \dots, \varepsilon_{cN_c})$ とおき、
 $F(\vec{\varepsilon})$ を $\vec{\varepsilon}$ の累積分布関数とする。

(1)はこのように書き直せる

$$P_{cn} = \int_{\varepsilon_{cn} = -\infty}^{+\infty} F_{cn}(\langle V_{cn} + \varepsilon_{cn} - V_{bm} \rangle) d\varepsilon_{cn} \quad (2)$$

F_{cn} …… F の cn に関する微分
 $\langle V_{cn} + \varepsilon_{cn} - V_{bm} \rangle$ …… bm で添字された成分を持つベクトル

注：式(1)→式(2)

$$\begin{aligned}
 P_{cn} &= \text{Prob}[U_{cn} > U_{bm} && \text{for } bm \neq cn] && (1) \\
 &= \text{Prob}[V_{cn} + \varepsilon_{cn} > V_{bm} + \varepsilon_{bm} && \text{for } bm \neq cn] \\
 &= \text{Prob}[V_{cn} + \varepsilon_{cn} - V_{bm} > \varepsilon_{bm} && \text{for } bm \neq cn]
 \end{aligned}$$

一方、

$F(\vec{\varepsilon}) \cdots \vec{\varepsilon}$ の累積分布関数

$F_{cn}(\vec{\varepsilon}) \cdots \vec{\varepsilon}$ の累積分布関数をcnで微分したもの

$\stackrel{\text{def}}{=} \varepsilon_{cn}$ の確率密度と ε_{bm} の累積分布を表す関数

$$F_{cn}(\langle V_{cn} + \varepsilon_{cn} - V_{bm} \rangle) \cdot \Delta \varepsilon_{cn} \quad (2')$$

$\stackrel{\text{def}}{=} \text{cn}$ 以外の要素 bm について全て $\langle V_{cn} + \varepsilon_{cn} - V_{bm} \rangle > \varepsilon_{bm}$ であり、かつ ε_{cn} が $\Delta \varepsilon_{cn}$ の幅に入っている確率

$$\int_{\varepsilon_{cn}=-\infty}^{+\infty} F_{cn}(\langle V_{cn} + \varepsilon_{cn} - V_{bm} \rangle) d\varepsilon_{cn}$$

$\stackrel{\text{def}}{=} (2')$ を $-\infty < \varepsilon_{cn} < +\infty$ まで積分した値。つまりcn以外の要素 bm について全て $\langle V_{cn} + \varepsilon_{cn} - V_{bm} \rangle > \varepsilon_{bm}$ である確率。 よって(2)が導ける。

式(2)の確率分布うち、よく使われるモデルがMNLモデルである。 ε_{cn} は独立して分布し、その分布をガンベル型の極値分布

$$Prob(\varepsilon_{cn} < \varepsilon) = \exp(-e^{-\varepsilon}) \quad (4)$$

とすると、以下の式が得られる。

$$P_{cn} = \exp(V_{cn}) / \sum_{b=1}^C \sum_{m=1}^{N_b} \exp(V_{bm}) \quad (3)$$

ただしこのモデルはIIA特性によって成り立っており、 ε_{cn} が独立と言えない場合には非現実的である。

ガンベル型の極値分布

累積分布関数 $F(\varepsilon) = \exp[-\exp\{-\mu(\varepsilon - \eta)\}]$

確率密度関数 $f(\varepsilon) = \mu \exp\{-\mu(\varepsilon - \eta)\} \exp[\exp\{-\mu(\varepsilon - \eta)\}]$

μ : 分布のばらつきを表すスケールパラメータ

η : 分布の位置を表すロケーションパラメータ

ロジットモデルでは効用確定項の大小と誤差項、スケールパラメータ μ に注目しているので、 $\eta = 0$ として扱うことが一般的
また、式を簡潔にするためにスケールパラメータ $\mu = 1$ とすることが多い

今回は、 $\eta = 0$, $\mu = 1$ とし、式(4)の $Prob(\varepsilon_{cn} < \varepsilon) = \exp(-e^{-\varepsilon})$ が出る。

ガンベル分布の性質

①最頻値は η , 平均値は $\eta + \frac{\gamma}{\mu}$ (γ :オイラ一定数), 分散は $\frac{\pi^2}{6\mu^2}$

② $\varepsilon_1, \varepsilon_2$ がそれぞれ $(\eta_1, \mu), (\eta_2, \mu)$ のパラメータを持つ互いに独立なガンベル分布に従うとき、 $\varepsilon = \varepsilon_1 - \varepsilon_2$ は次のロジスティック分布に従う。

$$F(\varepsilon) = \frac{1}{1 + \exp\{\mu(\eta_2 - \eta_1 - \varepsilon)\}}$$

③ $\varepsilon_1, \varepsilon_2 \dots \varepsilon_J$ がそれぞれ $(\eta_1, \mu), (\eta_2, \mu) \dots (\eta_J, \mu)$ のパラメータを持つ互いに独立なガンベル分布に従うとき、 $\max.(\varepsilon_1, \varepsilon_2 \dots \varepsilon_J)$ もガンベル分布に従い、そのパラメータは次のようになる。

$$\left(\frac{1}{\mu} \ln \sum_{j=1}^J \exp(\mu\eta_j), \mu \right)$$

注：式(4)→式(3)

式(4)の $Prob(\varepsilon_{cn} < \varepsilon) = \exp(-e^{-\varepsilon})$ より

$$\begin{aligned} Prob(\varepsilon_{bm} < V_{cn} + \varepsilon_{cn} - V_{bm}) &= \exp(-e^{-V_{cn} - \varepsilon_{cn} + V_{bm}}) \\ &= \exp(-e^{-\varepsilon_{cn}} \times e^{V_{bm} - V_{cn}}) \end{aligned}$$

$$\begin{aligned} P_{cn} &= \int_{\varepsilon_{cn}=-\infty}^{+\infty} F_{cn}(\langle V_{cn} + \varepsilon_{cn} - V_{bm} \rangle) d\varepsilon_{cn} \\ &= \int_{\varepsilon_{cn}=-\infty}^{+\infty} f(\varepsilon_{cn}) \prod_{bm \neq cn} Prob(\varepsilon_{bm} < V_{cn} + \varepsilon_{cn} - V_{bm}) d\varepsilon_{cn} \\ &= \int_{\varepsilon_{cn}=-\infty}^{+\infty} \exp(-\varepsilon_{cn}) \exp\{-e^{-\varepsilon_{cn}}\} \prod_{bm \neq cn} \exp(-e^{-\varepsilon_{cn}} \times e^{V_{bm} - V_{cn}}) d\varepsilon_{cn} \\ &= \int_{\varepsilon_{cn}=-\infty}^{+\infty} e^{-\varepsilon_{cn}} \exp\{-e^{-\varepsilon_{cn}} \times e^{V_{cn} - V_{cn}}\} \exp\left\{ \sum_{bm \neq cn} (-e^{-\varepsilon_{cn}} \times e^{V_{bm} - V_{cn}}) \right\} d\varepsilon_{cn} \\ &= \int_{\varepsilon_{cn}=-\infty}^{+\infty} e^{-\varepsilon_{cn}} \exp\left\{ \sum_{b=1}^C \sum_{m=1}^{N_b} (-e^{-\varepsilon_{cn}} \times e^{V_{bm} - V_{cn}}) \right\} d\varepsilon_{cn} \end{aligned}$$

注：式(4)→式(3)

$$\begin{aligned}
 &= \int_{\varepsilon_{cn}=-\infty}^{+\infty} e^{-\varepsilon_{cn}} \exp \left\{ \sum_{b=1}^C \sum_{m=1}^{N_b} (-e^{-\varepsilon_{cn}} \times e^{V_{bm}-V_{cn}}) \right\} d\varepsilon_{cn} \\
 &= \int_{\varepsilon_{cn}=-\infty}^{+\infty} e^{-\varepsilon_{cn}} \exp \left\{ -e^{-\varepsilon_{cn}} \sum_{b=1}^C \sum_{m=1}^{N_b} e^{V_{bm}-V_{cn}} \right\} d\varepsilon_{cn} \\
 &= \frac{1}{\sum_{b=1}^C \sum_{m=1}^{N_b} \exp(V_{bm} - V_{cn})} \left[\exp \left\{ -e^{-\varepsilon_{cn}} \sum_{b=1}^C \sum_{m=1}^{N_b} e^{V_{bm}-V_{cn}} \right\} \right]_{\varepsilon_{cn}=-\infty}^{+\infty} \\
 &= \frac{1}{\sum_{b=1}^C \sum_{m=1}^{N_b} \exp(V_{bm} - V_{cn})} \\
 &= \exp(V_{cn}) / \sum_{b=1}^C \sum_{m=1}^{N_b} \exp(V_{bm})
 \end{aligned}$$

よって式(3)が導出できた。

効用関数の確定項 V_{cn} を、線形のパラメーターで表現する。

$$V_{cn} = \beta' x_{cn} + \alpha' y_c \quad (5)$$

$x_{cn} \stackrel{\text{def}}{=} \text{コミュニティと住戸の両方で変化する属性(例：職場へのアクセス)のベクトル}$

$y_c \stackrel{\text{def}}{=} \text{コミュニティのみで変化する属性(例：コミュニティのレクリエーション施設の利用可能性)のベクトル}$

$\alpha, \beta \stackrel{\text{def}}{=} \text{未知パラメータのベクトル。右肩のダッシュは転置を表す。}$

$$x_{cn} = \begin{pmatrix} x_{cn;1} \\ \vdots \\ x_{cn;l} \end{pmatrix}, y_c = \begin{pmatrix} y_{c;1} \\ \vdots \\ y_{c;k} \end{pmatrix}, \alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_l \end{pmatrix}$$

$$\begin{aligned} V_{cn} = \beta' x_{cn} + \alpha' y_c &= (\beta_1 \quad \cdots \quad \beta_l) \begin{pmatrix} x_{cn;1} \\ \vdots \\ x_{cn;l} \end{pmatrix} + (\alpha_1 \quad \cdots \quad \alpha_k) \begin{pmatrix} y_{c;1} \\ \vdots \\ y_{c;k} \end{pmatrix} \\ &= \beta_1 x_{cn;1} + \cdots + \beta_l x_{cn;l} + \alpha_1 y_{c;1} + \cdots + \alpha_k y_{c;k} \end{aligned}$$

式(3)のMNLモデルを、コミュニティが与えられた場合の条件付き選択確率 $P_{n|c}$ と、コミュニティの選択確率 P_c に分けて書き換える。

$$\begin{aligned} P_{n|c} &= \exp(V_{cn}) / \sum_{m=1}^{N_c} \exp(V_{cm}) = \exp(\beta' x_{cn} + \alpha' y_c) / \sum_{m=1}^{N_c} \exp(\beta' x_{cm} + \alpha' y_c) \\ &= \exp(\beta' x_{cn}) / \sum_{m=1}^{N_c} \exp(\beta' x_{cm}) \quad (6) \end{aligned}$$

$$\begin{aligned} P_c &= \sum_{n=1}^{N_c} \exp(V_{cn}) / \sum_{b=1}^C \sum_{m=1}^{N_b} \exp(V_{bm}) = \sum_{n=1}^{N_c} \exp(\beta' x_{cn} + \alpha' y_c) / \sum_{b=1}^C \sum_{m=1}^{N_b} \exp(\beta' x_{bm} + \alpha' y_b) \\ &= \exp(\alpha' y_c) \left[\sum_{m=1}^{N_c} \exp(\beta' x_{cm}) \right] / \sum_{b=1}^C \exp(\alpha' y_b) \left[\sum_{m=1}^{N_b} \exp(\beta' x_{bm}) \right] \quad (7) \end{aligned}$$

$$I_c = \log \left[\sum_{n=1}^{N_c} \exp(\beta' x_{cn}) \right] \quad (8)$$

を定義すると、(6),(7)は次のように書ける

$$P_{n|c} = \exp(\beta' x_{cn}) / \exp(I_c) \quad (9)$$

$$P_c = \exp(\alpha' y_c + I_c) / \sum_{b=1}^C \exp(\alpha' y_b + I_b) \quad (10)$$

ここで、

式(6)からパラメータ β を推定

↓

I_c を定義

↓

式(10)から α を推定

というモデル推定方法が考えられる

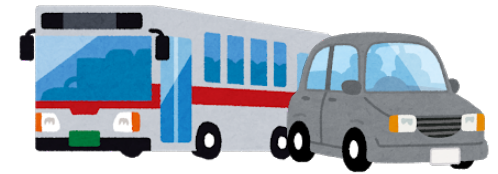
逐次的で効率が落ちるが、推定の各段階で考慮する選択肢やパラメータを節約

4. NLモデル(ネステッドロジットモデル) 16

例) 赤バス青バス問題

選択肢が赤バスと自動車だけの場合、選択確率は

赤バス:自動車=0.5:0.5



とします。では、選択肢に青バスが加わったら？

MNLモデルでは

赤バス:青バス:自動車
=0.33:0.33:0.33



0.25:0.25:0.5では？ ≡ 選択肢間に相関があるのでは？

→NLモデル(Nested Logit Model)

4. NLモデル(ネステッドロジットモデル)

NLモデル(Nested Logit Model)
ロジットモデルのIIA特性を緩和

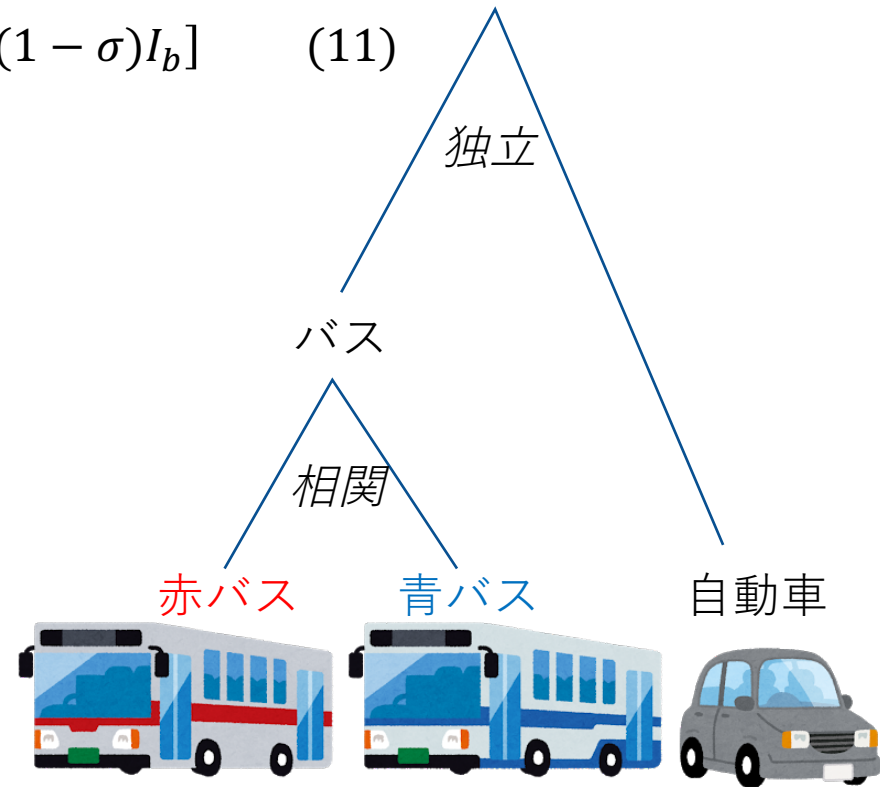
$$P_c = \exp[\alpha' y_c + (1 - \sigma)I_c] / \sum_{b=1}^c \exp[\alpha' y_b + (1 - \sigma)I_b] \quad (11)$$

$0 \leq \sigma \leq 1$: 相関の度合いを表すパラメータ

$\sigma = 0$ の時、NLモデルはMNLモデルとなる
→MNLモデルを一般化したもの

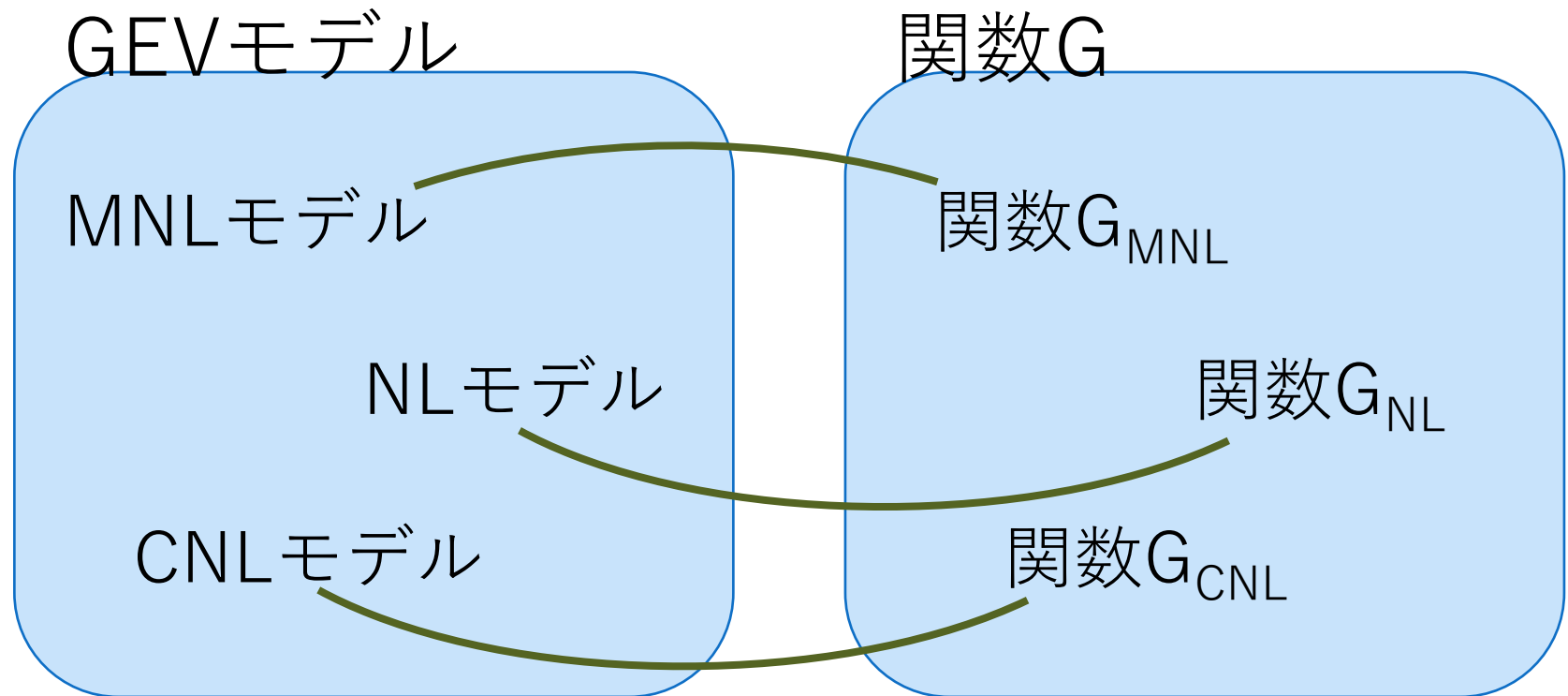
$$P_{cn} = P_n | P_c$$

の P_c に(11)を代入した形となる。



GEVモデル(Generalized Extreme Value Model)

ある条件を満たす関数Gによって特徴づけられるモデル
属性間の依存性の一般的なパターンを可能にする
分析的に扱いやすいクローズドフォーム
MNLモデル、NLモデルなどもGEVモデルの一種



G関数の特徴付ける、満たすべき条件

1. $G(y_1, \dots, y_J) \geq 0$ for $(\forall j \ y_j \geq 0)$

2. Gは1次の同次関数

$$G(\rho y_1, \dots, \rho y_J) = \rho G(y_1, \dots, y_J)$$

3. $\forall j \ \lim_{y_j \rightarrow \infty} G = +\infty$

4. $\frac{\partial^k G}{\partial y_{i_1} \dots \partial y_{i_k}} \begin{cases} \geq 0 & (\text{if } k \text{ is odd}) \\ \leq 0 & (\text{if } k \text{ is even}) \end{cases}$

この条件下で次のようになるものをGEVモデルとする。

$$P_i = y_i G_i / G \quad (12)$$

文字の定義

$$y_1 = \exp(V_1)$$

$$G = G(y_1, \dots, y_J)$$

$$G_i = \partial G / \partial y_i$$

MNLモデルとGEVモデル

MNLモデルにおける選択確率

$$P_i = \exp(V_i) / \sum_{j=1}^J \exp(V_j)$$

この部分をGとおく

$$G = \sum_{j=1}^J y_j$$

関数Gが満たすべき条件

1. $G(y_1, \dots, y_J) \geq 0$ for $(\forall j \ y_j \geq 0)$
2. $G(\rho y_1, \dots, \rho y_J) = \rho G(y_1, \dots, y_J)$
3. $\forall j \ \lim_{y_j \rightarrow \infty} G = \infty$
4. $\frac{\partial^k G}{\partial y_{i_1} \dots \partial y_{i_k}} \begin{cases} \geq 0 \text{ (if } k \text{ is odd)} \\ \leq 0 \text{ (if } k \text{ is even)} \end{cases}$

4つの条件の確認

1. OK
2. $\sum_{j=1}^J \rho y_j = \rho \sum_{j=1}^J y_j$ OK
3. OK
4. $\partial G / \partial y_i = 1 \geq 0$
2階以上の偏導関数の値は全て0
→OK

式(12) $P_i = y_i G_i / G$ を用いて P_i を求めると

$$P_i = y_i G_i / G = y_i / \sum_{j=1}^J y_j = \exp(V_i) / \sum_{j=1}^J \exp(V_j)$$

もう少し一般化し

$$F(\varepsilon_1, \dots, \varepsilon_J) = \exp\{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$$

とすると、式(12) $P_i = y_i G_i / G$ が導ける。

証明

$$\text{式(2)より } P_i = \int_{\varepsilon_i=-\infty}^{+\infty} F_i(\langle V_i + \varepsilon_i - V_j \rangle) d\varepsilon_i$$

$$= \int_{\varepsilon_i=-\infty}^{+\infty} \frac{e^{-\varepsilon_i} G_i(e^{-V_i-\varepsilon_i+V_1}, \dots, e^{-V_i-\varepsilon_i+V_J})}{\text{関数 } G \text{ の条件2より、定数関数}} \cdot \exp\{-G(e^{-V_i-\varepsilon_i+V_1}, \dots, e^{-V_i-\varepsilon_i+V_J})\} d\varepsilon_i$$

$$= \int_{\varepsilon_i=-\infty}^{+\infty} \frac{e^{-\varepsilon_i} G_i(e^{V_1}, \dots, e^{V_J})}{\text{関数 } G \text{ の条件2より、定数関数}} \cdot \exp\{-e^{-V_i} e^{-\varepsilon_i} G(e^{V_1}, \dots, e^{V_J})\} d\varepsilon_i$$

$$= e^{V_i} \frac{G_i(e^{V_1}, \dots, e^{V_J})}{G(e^{V_1}, \dots, e^{V_J})}$$

$$= \frac{y_i G_i}{G}$$

ただし本当は

$$F(\varepsilon_1, \dots, \varepsilon_J) = \exp\{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$$

で表される $\exp\{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$ が多変量極値分布の累積分布関数

$$\Leftrightarrow \begin{cases} 1) \lim_{\varepsilon_j \rightarrow -\infty} F = 0 \text{ かつ } \lim_{\{\varepsilon_j\} \rightarrow +\infty} F = 1 \\ 2) F \text{ が単調増加かつ連続、つまり偏微分が非負であること} \end{cases}$$

を示す必要がある

→証明は2015年理論勉強会#3行動モデルを見てください。

1) $\varepsilon_j \rightarrow -\infty$ ならば $e^{-\varepsilon_j} \rightarrow +\infty$

関数Gの条件3 $\forall j \lim_{y_j \rightarrow \infty} G = +\infty$ より $\lim_{\varepsilon_j \rightarrow -\infty} G = +\infty$ then $\lim_{\varepsilon_j \rightarrow -\infty} F = 0$

関数Gの条件2一次同次性より $\lim_{\{\varepsilon_j\} \rightarrow +\infty} G = 0$ then $\lim_{\varepsilon_j \rightarrow +\infty} F = 1$

2) $F(\varepsilon_1, \dots, \varepsilon_J) = \exp\{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$ の偏微分が非負であることを示す

(2) F が単調増加かつ連続である、つまり F の偏微分が非負である。

これを示す。

以下のように Q_k を定める。

$$\begin{cases} Q_1 = G_1 \\ Q_k = Q_{k-1}G_k - \partial Q_{k-1}/\partial Y_k \end{cases}$$

$k = 1, \dots, J$ について $Q_k \geq 0$ であることを帰納法で示す。

i) $k = 1$ のとき

G の満たす性質4より

$$G_1 = \partial G / \partial Y_1 \geq 0$$

$$\text{よって } Q_1 = G_1 \geq 0$$

ii) $k \geq 2$ のとき

$Q_{k-1} \geq 0$ と仮定する。

$$\begin{cases} Q_1 = G_1 \\ Q_k = Q_{k-1}G_k - \partial Q_{k-1}/\partial Y_k \end{cases} \text{より}$$

Q_k は G のさまざまな階数の偏微分からなる項の和。

G の満たす性質4より、

Q_k を Y_j で偏微分する

$\rightarrow Q_k$ を構成する G の n 階の偏微分が $(n+1)$ 階の偏微分になる

$\rightarrow Q_k$ の各項の正負が入れ替わる

$$\rightarrow Q_{k-1} \geq 0 \text{ なら } \partial Q_{k-1}/\partial Y_k \leq 0$$

$$\text{また, } G_k = \partial G / \partial Y_k \geq 0 \text{ より } Q_{k-1}G_k \geq 0$$

$$\text{よってこのとき, } Q_k \geq 0$$

i), ii) より、 $k = 1, \dots, J$ について $Q_k \geq 0$

$$F(\varepsilon_1, \dots, \varepsilon_J) = \exp\{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$$

$k = 1, \dots, J$ について

$$\partial^k F / \partial \varepsilon_1, \dots, \partial \varepsilon_k \geq 0$$

を帰納法により示す。

i) $k = 1$ のとき

$$\frac{\partial F}{\partial \varepsilon_1} = \exp\{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\} \cdot \frac{\partial}{\partial \varepsilon_1} \{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$$

$$= F \cdot \frac{\partial e^{-\varepsilon_1}}{\partial \varepsilon_1} \cdot \frac{\partial}{\partial e^{-\varepsilon_1}} \{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$$

$$= F \cdot (-e^{-\varepsilon_1}) \cdot (-G_1)$$

$$= e^{-\varepsilon_1} Q_1 F \geq 0$$

これと同様に考えて、

$$\frac{\partial F}{\partial \varepsilon_j} = e^{-\varepsilon_j} G_j F$$

ii) $k \geq 2$ のとき

$$\frac{\partial^{k-1} F}{\partial \varepsilon_1, \dots, \partial \varepsilon_{k-1}} = e^{-\varepsilon_1} \dots e^{-\varepsilon_{k-1}} Q_{k-1} F$$

と仮定する。このとき

$$\begin{aligned} \frac{\partial^k F}{\partial \varepsilon_1, \dots, \partial \varepsilon_k} &= e^{-\varepsilon_1} \dots e^{-\varepsilon_{k-1}} \cdot \frac{\partial}{\partial \varepsilon_k} (Q_{k-1} F) \\ &= e^{-\varepsilon_1} \dots e^{-\varepsilon_{k-1}} \left(\frac{\partial Q_{k-1}}{\partial \varepsilon_k} \cdot F + Q_{k-1} \cdot \frac{\partial F}{\partial \varepsilon_k} \right) \\ &= \frac{\partial e^{-\varepsilon_k}}{\partial \varepsilon_k} \cdot \frac{\partial Q_{k-1}}{\partial e^{-\varepsilon_k}} \cdot F + Q_{k-1} e^{-\varepsilon_k} G_k F \\ &= (-e^{-\varepsilon_k} \cdot \frac{\partial Q_{k-1}}{\partial Y_k} + Q_{k-1} e^{-\varepsilon_k} G_k) F \\ &= e^{-\varepsilon_k} \left(Q_{k-1} G_k - \frac{\partial Q_{k-1}}{\partial Y_k} \right) F \\ &= e^{-\varepsilon_k} Q_k F \\ &= e^{-\varepsilon_1} \dots e^{-\varepsilon_k} Q_k F \end{aligned}$$

i), ii) より

$k = 1, \dots, J$ について

$$\partial^k F / \partial \varepsilon_1, \dots, \partial \varepsilon_k \geq 0$$

$k = 1, \dots, J$ について

$$\partial^k F / \partial \varepsilon_1, \dots, \partial \varepsilon_k \geq 0$$

つまり、

(2) F が単調増加かつ連続である、つまり F の偏微分が非負である。

を示せた。

$\rightarrow F$ は累積分布関数ということが示せた！！

ここで、 i 以外の j について $\varepsilon_j = +\infty$ を与えたとき

$$F = \exp\{-a_i e^{-\varepsilon_i}\}, \quad a_i = G(0, \dots, 0, 1, 0, \dots, 0) \quad \text{番目}$$

\rightarrow ガンベル分布(一変量極値分布)に一致！

$\rightarrow F$ は多変量極値分布である

以上より

$$F(\varepsilon_1, \dots, \varepsilon_J) = \exp\{-G(e^{-\varepsilon_1}, \dots, e^{-\varepsilon_J})\}$$

は多変量極値分布の累積分布関数である

最大効用 $\max_j U_j$ の期待値を \bar{U} とし、以下のように定義する。

$$\bar{U} = \int_{\vec{\varepsilon}=-\infty}^{+\infty} \max_i (V_i + \varepsilon_i) f(\vec{\varepsilon}) d\vec{\varepsilon} \quad (13)$$

また、これは以下の式を満たす。

$$\bar{U} = \log G[\exp(V_1), \dots, \exp(V_J)] + \gamma \quad (14)$$

ただし、 $\gamma \approx 0.57721$ はオイラーの定数を表す。

また、 $\bar{U} = \sum_{j=1}^J P_j (V_j + \varepsilon_j)$ より、

$$\begin{aligned} \partial \bar{U} / \partial V_i &= P_i \\ P_i &= \partial \bar{U} / \partial V_i \end{aligned} \quad (15)$$

注)式(14) $\bar{U} = \log G[\exp(V_1), \dots, \exp(V_J)] + \gamma$ の証明

効用関数 $U_i = V_i + \varepsilon_i$ の最大値を累積分布 F を用いて考える。

$$\begin{aligned}
 \text{Prob}\left(\max_i U_i < U\right) &= \text{Prob}(U_i < U \text{ for } \forall i) \\
 &= \text{Prob}(\varepsilon_i < U - V_i \text{ for } \forall i) \\
 &= F(\langle U - V_i \rangle) \\
 &= \exp\left\{-G\left(e^{V_1-U}, \dots, e^{V_J-U}\right)\right\} \\
 &= \exp\left\{-\exp(-U) \cdot G\left(e^{V_1}, \dots, e^{V_J}\right)\right\} \\
 &= \exp\left\{-\exp(-U) \cdot \exp(\log G)\right\} \\
 &= \exp\left\{-\exp(-U) \cdot \exp(\log G)\right\} \\
 &= \exp\left[-\exp\left\{-(U - \log G)\right\}\right]
 \end{aligned}$$

よって、 $\max_i U_i$ の分布はスケールパラメータ $\mu = 1$,

ロケーションパラメータ $\eta = \log G$ のガンベル分布

注)式(14) $\bar{U} = \log G[\exp(V_1), \dots, \exp(V_J)] + \gamma$ の証明
次に、 $\max_i U_i$ の期待値を求める。 $\max_i U_i$ を U とおく。

$$\bar{U} = \int_{\vec{\varepsilon}=-\infty}^{+\infty} \max_i (V_i + \varepsilon_i) f(\vec{\varepsilon}) d\vec{\varepsilon}$$

$$\bar{U} = \int_{U=-\infty}^{+\infty} \frac{U \exp\{-(U - \log G)\} \exp[-\exp\{-(U - \log G)\}]}{U_i \text{の最大値 } U \text{の分布。ガンベル分布に従う。}} dU$$

$z = \exp\{-(U - \log G)\}$ とおくと、 $U = \log G - \log z$ より、 $dU = -\frac{1}{z} dz$

$$(\text{上式}) = - \int_{z=0}^{+\infty} (\log G - \log z) z e^{-z} \left(-\frac{1}{z}\right) dz$$

$$= \int_{z=0}^{+\infty} (\log G - \log z) e^{-z} dz$$

$$= \log G \int_{z=0}^{+\infty} e^{-z} dz - \int_{z=0}^{+\infty} e^{-z} \log z dz$$

$$= \log G + \gamma$$

オイラーの定数の定義
 $\gamma = - \int_{z=0}^{+\infty} e^{-z} \log z dz$
より

以上より $\bar{U} = \log G[\exp(V_1), \dots, \exp(V_J)] + \gamma$ が証明された

注)式(15)から式(12)を導くこともできる。

$$P_i = \partial \bar{U} / \partial V_i \quad (15)$$

式(14) $\bar{U} = \log G[\exp(V_1), \dots, \exp(V_J)] + \gamma$ より

$$\begin{aligned} P_i &= \partial \bar{U} / \partial V_i \\ &= \frac{\partial \bar{U}}{\partial G} \cdot \frac{\partial G}{\partial y_i} \cdot \frac{\partial y_i}{\partial V_i} \\ &= \frac{1}{G} \cdot G_i \cdot \{\exp(V_i)\}' \\ &= \frac{1}{G} \cdot G_i \cdot \exp(V_i) \\ &= \frac{y_i G_i}{G} \end{aligned}$$

条件を満たす一般的なG関数の例は以下の通り。

$$G(y) = \sum_{m=1}^M a_m \left[\sum_{i \in B_m} y_i^{1/(1-\sigma_m)} \right]^{1-\sigma_m} \quad (16)$$

ここで、 $B_m \subseteq \{1, \dots, J\}$, $\cup_{m=1}^M B_m = \{1, \dots, J\}$, $a_m > 0$, $0 \leq \sigma_m < 1$

例)同一クラスmに属する二選択肢の場合、次のようにかける。

$$G(y) = [y_1^{1/(1-\sigma)} + y_2^{1/(1-\sigma)}]^{1-\sigma} \quad (17)$$

ここで、この二変量極値分布において σ は二変量間の**相関**を表す。

式(16) $G(y) = \sum_{m=1}^M a_m [\sum_{i \in B_m} y_i^{1/(1-\sigma_m)}]^{1-\sigma_m}$ で表される選択確率は以下の式を満たす。

$$P_i = \sum_{m=1}^M P(i|B_m)P(B_m) \quad (18)$$

ただし

$$\begin{cases} P(i|B_m) = \exp[V_i/(1-\sigma_m)] / \sum_{j \in B_m} \exp[V_j/(1-\sigma_m)] & (\text{if } i \in B_m) \quad (19) \\ P(i|B_m) = 0 & (\text{if } i \notin B_m) \quad (20) \end{cases}$$

$$P(B_m) = \frac{a_m \left\{ \sum_{j \in B_m} \exp[V_j/(1-\sigma_m)] \right\}^{1-\sigma_m}}{\sum_{n=1}^M a_n \left\{ \sum_{k \in B_n} \exp[V_k/(1-\sigma_n)] \right\}^{1-\sigma_n}} \quad (21)$$

式(18)~(21)の証明

式(18)~(21)より $P_i = \sum_{m=1}^M P(i|B_m)P(B_m)$ (18)

$$\begin{aligned}
 P_i &= \sum_{\forall B_m \ni i} \frac{\exp[V_i/(1-\sigma_m)]}{\sum_{j \in B_m} \exp[V_j/(1-\sigma_m)]} \frac{a_m \left\{ \sum_{j \in B_m} \exp[V_j/(1-\sigma_m)] \right\}^{1-\sigma_m}}{\sum_{n=1}^M a_n \left\{ \sum_{k \in B_n} \exp[V_k/(1-\sigma_n)] \right\}^{1-\sigma_n}} \\
 &= \sum_{\forall B_m \ni i} \frac{a_m \exp[V_i/(1-\sigma_m)] \left\{ \sum_{j \in B_m} \exp[V_j/(1-\sigma_m)] \right\}^{-\sigma_m}}{\sum_{n=1}^M a_n \left\{ \sum_{k \in B_n} \exp[V_k/(1-\sigma_n)] \right\}^{1-\sigma_n}}
 \end{aligned}$$

一方、式(16)より

$$\begin{aligned}
 y_i G_i / G &= \frac{y_i \sum_{\forall B_m \ni i} a_m y_i^{\sigma_m/(1-\sigma_m)} \left[\sum_{i \in B_m} y_i^{1/(1-\sigma_m)} \right]^{-\sigma_m}}{\sum_{m=1}^M a_m \left[\sum_{i \in B_m} y_i^{1/(1-\sigma_m)} \right]^{1-\sigma_m}} \\
 &= \frac{\sum_{\forall B_m \ni i} \left\{ a_m y_i^{1/(1-\sigma_m)} \left[\sum_{i \in B_m} y_i^{1/(1-\sigma_m)} \right]^{-\sigma_m} \right\}}{\sum_{m=1}^M a_m \left[\sum_{i \in B_m} y_i^{1/(1-\sigma_m)} \right]^{1-\sigma_m}}
 \end{aligned}$$

よって一致するため、式(12) $P_i = y_i G_i / G$ が成り立つ。これにより式(18)~(21)は式(16), (12)を満たす。

GEVモデルと赤バス青バス問題

式(16) $G(y) = \sum_{m=1}^M a_m [\sum_{i \in B_m} y_i^{1/(1-\sigma_m)}]^{1-\sigma_m}$ において、二つのグループ、3つの選択肢 $B_{\text{車}} = \{\text{車}\}$, $B_{\text{バス}} = \{\text{赤}, \text{青}\}$ がある場合について考える。関数Gを以下のように仮定する。

$$G(y_{\text{車}}, y_{\text{赤}}, y_{\text{青}}) = y_{\text{車}} + \left[y_{\text{赤}}^{1/(1-\sigma)} + y_{\text{青}}^{1/(1-\sigma)} \right]^{1-\sigma} \quad (22)$$

この時、各選択確率は次のようになる

$P(\text{車} | \text{車}, \text{赤}, \text{青})$

$$= \exp(V_{\text{車}}) / \left(\exp(V_{\text{車}}) + \left\{ \exp[V_{\text{赤}}/(1-\sigma)] + \exp[V_{\text{青}}/(1-\sigma)] \right\}^{1-\sigma} \right) \quad (23)$$

$P(\text{赤} | \text{車}, \text{赤}, \text{青})$

$$= \frac{\exp[V_{\text{赤}}/(1-\sigma)] \left\{ \exp[V_{\text{赤}}/(1-\sigma)] + \exp[V_{\text{青}}/(1-\sigma)] \right\}^{-\sigma}}{\exp(V_{\text{車}}) + \left\{ \exp[V_{\text{赤}}/(1-\sigma)] + \exp[V_{\text{青}}/(1-\sigma)] \right\}^{1-\sigma}} \quad (24)$$

GEVモデルと赤バス青バス問題

ここで、選択肢が車と赤バスしかない場合、選択確率は $V_{\text{青}} \rightarrow -\infty$ を代入し、次のような二項式となる。

$$P(\text{車}|\text{車}, \text{赤}) = \exp(V_{\text{車}}) / (\exp(V_{\text{車}}) + \exp(V_{\text{赤}})) \quad (25)$$

一方、選択肢が赤バスと青バスしかない場合、選択確率は式(17)の二項ロジット形式となる。

$$\begin{aligned} & P(\text{赤}|\text{赤}, \text{青}) \\ &= \exp[V_{\text{赤}}/(1-\sigma)] / (\exp[V_{\text{赤}}/(1-\sigma)] + \exp[V_{\text{青}}/(1-\sigma)]) \end{aligned} \quad (26)$$

GEVモデルと赤バス青バス問題

次に、相関の度合いを表す σ について考える。

$\sigma = 0 \rightarrow$ MNLモデルとなる。

極限 $\sigma \rightarrow 1$ の場合、以下の確率となる。

$$\begin{aligned}
 &P(\text{車}|\text{車}, \text{赤}, \text{青}) \\
 &= \exp(V_{\text{車}}) / \left(\exp(V_{\text{車}}) + \max[\exp(V_{\text{赤}}), \exp(V_{\text{青}})] \right) \quad (27)
 \end{aligned}$$

$$\begin{aligned}
 &P(\text{赤}|\text{車}, \text{赤}, \text{青}) \\
 &= \begin{cases} \exp(V_{\text{赤}}) / \left(\exp(V_{\text{車}}) + \exp(V_{\text{赤}}) \right) & (\text{if } V_{\text{赤}} > V_{\text{青}}) \\ \exp(V_{\text{赤}}) / 2 \left(\exp(V_{\text{車}}) + \exp(V_{\text{赤}}) \right) & (\text{if } V_{\text{赤}} = V_{\text{青}}) \\ 0 & (\text{if } V_{\text{赤}} < V_{\text{青}}) \end{cases} \quad (28)
 \end{aligned}$$

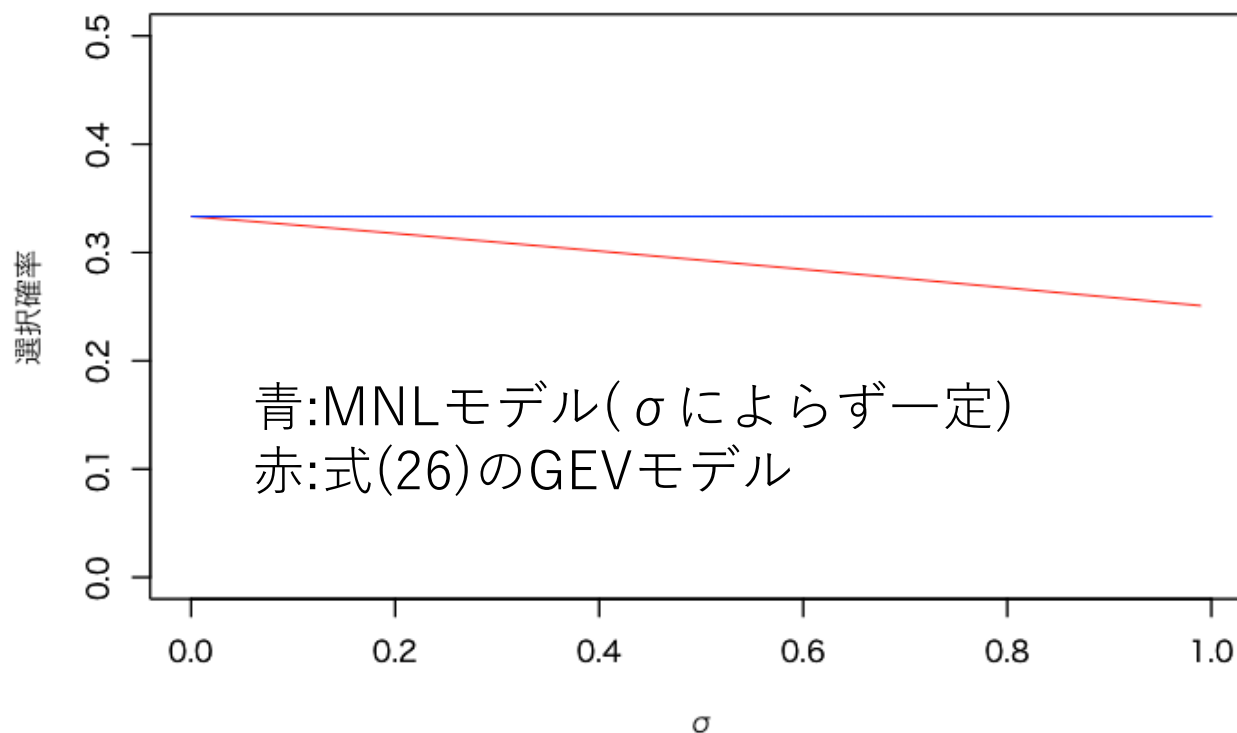
この極端な場合、車に対しバスを一つの選択肢として扱うことになる。

GEVモデルと赤バス青バス問題

σ と選択確率との関係

$V_{車} = V_{赤} = V_{青} = 1$ と仮定

σ と選択確率



GEVモデル式(18) $P_i = \sum_{m=1}^M P(i|B_m)P(B_m)$ の変数を以下のように仮定することでNLモデルの式(9), (11)を得ることができる。

選択肢番号 $i \Rightarrow$ コミュニティ番号 c と住戸番号 n による cn

類似性を持つグループ番号 $m \Rightarrow$ コミュニティ番号 c

選択肢集合 $B_m \Rightarrow B_c = \{c_1, \dots, c_{N_c}\}$

類似性を示す係数 $\sigma_m \Rightarrow$ 共通の値 σ を持つ

影響度の重み付けをする $a_m \Rightarrow$ 共通の値 1 を持つ

式(18)~(21)より、

$$P_i = \sum_{\forall B_m \ni i} \frac{a_m \exp[V_i/(1 - \sigma_m)] \left\{ \sum_{j \in B_m} \exp[V_j/(1 - \sigma_m)] \right\}^{-\sigma_m}}{\sum_{n=1}^M a_n \left\{ \sum_{k \in B_n} \exp[V_k/(1 - \sigma_n)] \right\}^{1 - \sigma_n}}$$

よって

$$P_{cn} = \frac{\exp[V_{cn}/(1 - \sigma)] \left\{ \sum_{m=1}^{N_c} \exp[V_{cm}/(1 - \sigma)] \right\}^{-\sigma}}{\sum_{b=1}^C \left\{ \sum_{m=1}^{N_b} \exp[V_{bm}/(1 - \sigma)] \right\}^{1 - \sigma}} \quad (29)$$

一方、 $J = \sum_{b=1}^C N_b$ 個の選択肢が C 個のコミュニティ $\{1, \dots, C\}$ に分けられていると考え、関数 G を以下のように定義して式(12) $P_i = y_i G_i / G$ を用いても式(29)を導くことができる。

$$G(y_{11}, \dots, y_{CN_C}) = \sum_{b=1}^C \left(\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)} \right)^{1-\sigma}$$

G の性質を持つ条件の確認

関数 G が満たすべき条件

1. $G(y_1, \dots, y_J) \geq 0$ for $(\forall j \ y_j \geq 0)$
2. $G(\rho y_1, \dots, \rho y_J) = \rho G(y_1, \dots, y_J)$
3. $\forall j \ \lim_{y_j \rightarrow \infty} G = \infty$
4. $\partial^k G / \partial y_{i_1} \dots \partial y_{i_k} \begin{cases} \geq 0 & (\text{if } k \text{ is odd}) \\ \leq 0 & (\text{if } k \text{ is even}) \end{cases}$

4つの条件の確認

1. OK
2. $\sum_{b=1}^C (\sum_{m=1}^{N_b} (\rho y_{bm})^{1/(1-\sigma)})^{1-\sigma}$
 $= \sum_{b=1}^C (\rho^{1/(1-\sigma)} \sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)})^{1-\sigma}$
 $= \sum_{b=1}^C (\rho^{1/(1-\sigma)})^{1-\sigma} (\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)})^{1-\sigma}$
 $= \rho \sum_{b=1}^C (\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)})^{1-\sigma}$ OK
3. OK

$$G(y_{11}, \dots, y_{CN_C}) = \sum_{b=1}^C \left(\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)} \right)^{1-\sigma}$$

が条件4. $\partial^k G / \partial y_{i_1} \dots \partial y_{i_k} \begin{cases} \geq 0 \text{ (if } k \text{ is odd)} \\ \leq 0 \text{ (if } k \text{ is even)} \end{cases}$ を満たすことの確認

コミュニティ c にある住戸 n について、

$$\begin{aligned} G_{cn} &= \frac{\partial}{\partial y_{cn}} \left(\sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right)^{1-\sigma} \\ &= (1-\sigma) \left(\sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right)^{-\sigma} \cdot \frac{1}{1-\sigma} y_{cn}^{\sigma/(1-\sigma)} \\ &= y_{cn}^{\sigma/(1-\sigma)} \left(\sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right)^{-\sigma} \geq 0 \end{aligned}$$

よって一階微分はOK

$$G(y_{11}, \dots, y_{CN_C}) = \sum_{b=1}^C \left(\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)} \right)^{1-\sigma}$$

が条件4. $\partial^k G / \partial y_{i_1} \dots \partial y_{i_k} \begin{cases} \geq 0 \text{ (if } k \text{ is odd)} \\ \leq 0 \text{ (if } k \text{ is even)} \end{cases}$ を満たすことの確認

コミュニティcにある住戸kについて、

$$\begin{aligned} G_{cn,ck} &= \frac{\partial}{\partial y_{ck}} y_{cn}^{\sigma/(1-\sigma)} \left(\sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right)^{-\sigma} \\ &= y_{cn}^{\sigma/(1-\sigma)} (-\sigma) \left(\sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right)^{-\sigma-1} \cdot \frac{1}{1-\sigma} y_{ck}^{\sigma/(1-\sigma)} \\ &= \frac{-\sigma}{1-\sigma} (y_{cn} y_{ck})^{\sigma/(1-\sigma)} \left(\sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right)^{-\sigma-1} \leq 0 \end{aligned}$$

住戸kがcに含まれればい場合は、 $G_{cn,ck} = 0$ より、まとめて $G_{cn,ck} \leq 0$ よって二階微分はOK、同様にこれより高階な偏微分でも成り立つ。

$$G(y_{11}, \dots, y_{CN_C}) = \sum_{b=1}^C \left(\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)} \right)^{1-\sigma}$$

より、式(12) $P_i = y_i G_i / G$ を用いても式(29)を導く。

$$\begin{aligned} P_{cn} &= \frac{y_{cn} \cdot y_{cn}^{-\sigma/(1-\sigma)} \cdot (1-\sigma) \cdot \frac{1}{(1-\sigma)} \left\{ \sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right\}^{-\sigma}}{\sum_{b=1}^C \left(\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)} \right)^{1-\sigma}} \\ &= \frac{y_{cn}^{1/(1-\sigma)} \left\{ \sum_{m=1}^{N_c} y_{cm}^{1/(1-\sigma)} \right\}^{-\sigma}}{\sum_{b=1}^C \left(\sum_{m=1}^{N_b} y_{bm}^{1/(1-\sigma)} \right)^{1-\sigma}} \\ &= \frac{\exp[V_{cn}/(1-\sigma)] \left\{ \sum_{m=1}^{N_c} \exp[V_{cm}/(1-\sigma)] \right\}^{-\sigma}}{\sum_{b=1}^C \left\{ \sum_{m=1}^{N_b} \exp[V_{bm}/(1-\sigma)] \right\}^{1-\sigma}} \end{aligned}$$

よって式(29)に一致

また式(29) $P_{cn} = \frac{\exp[V_{cn}/(1-\sigma)] \left\{ \sum_{m=1}^{N_c} \exp[V_{cm}/(1-\sigma)] \right\}^{-\sigma}}{\sum_{b=1}^C \left\{ \sum_{m=1}^{N_b} \exp[V_{bm}/(1-\sigma)] \right\}^{1-\sigma}}$ より、

$$P_c = \sum_{n=1}^{N_c} P_{cn} = \frac{\left\{ \sum_{m=1}^{N_c} \exp[V_{cm}/(1-\sigma)] \right\}^{1-\sigma}}{\sum_{b=1}^C \left\{ \sum_{m=1}^{N_b} \exp[V_{bm}/(1-\sigma)] \right\}^{1-\sigma}} \quad (30)$$

$$Pn|_c = P_{cn}/P_c = \exp[V_{cn}/(1-\sigma)] / \sum_{m=1}^{N_c} \exp[V_{cm}/(1-\sigma)] \quad (31)$$

一方で式(5) $V_{cn} = \beta' x_{cn} + \alpha' y_c$ より、

$$P_c = \frac{\{\sum_{m=1}^{N_c} \exp[V_{cm}/(1-\sigma)]\}^{1-\sigma}}{\sum_{b=1}^C \{\sum_{m=1}^{N_b} \exp[V_{bm}/(1-\sigma)]\}^{1-\sigma}} \quad (30) \text{より}$$

$$= \exp[\alpha' y_c + (1-\sigma)I_c] / \sum_{b=1}^C \exp[\alpha' y_b + (1-\sigma)I_b] \quad (32)$$

$$P_{n|c} = \exp[V_{cn}/(1-\sigma)] / \sum_{m=1}^{N_c} \exp[V_{cm}/(1-\sigma)] \quad (31) \text{より}$$

$$= \exp[\beta' x_{cn}/(1-\sigma)] / \sum_{m=1}^{N_c} \exp[\beta' x_{cn}/(1-\sigma)]$$

$$= \exp[\beta' x_{cn}/(1-\sigma)] / \exp(I_c) \quad (33)$$

ただし $I_c = \log \sum_{m=1}^{N_c} \exp[\beta' x_{cn}/(1-\sigma)]$ (34) とする。

よってNLモデルはGEVモデルの一部であり、以下のことがわかる。

- NLモデルがGEVモデルの確率的効用最大化と一致している。
- $(1 - \sigma)I_c$ は同じコミュニティ内の確率項の独立性を表すこと。

また、NLモデルで得られた I_c を用いて一般化し、GEVモデルを推定することができる。これは任意の深さのネストに拡張することができ、NLモデルがGEVモデルにおける確率的効用最大化と一致するための条件は I_c の係数 $(1 - \sigma)$ が0から1の間に入ることである。

式(3) $P_{cn} = \exp(V_{cn}) / \sum_{b=1}^C \sum_{m=1}^{N_b} \exp(V_{bm})$ で表されたMNLモデルは、利用可能な全選択肢からの選択を記述するのに適しており、データ収集と統計分析から実用的である。

→しかし実際には、全選択肢の分析は非現実的な量のデータ処理を必要とし、IIA条件が満たされない選択肢などで、MNLモデルの信憑性が歪みやすい。

→**現実的な選択肢数で適切なパラメータ θ^* を推定する方法を考える。**

まず、全選択肢の列挙はできないが、観測できるデータは全て独立な(IIA条件を満たす)場合を考える。

全選択肢の列挙はできないが、観測できるデータは全て独立な (IIA条件を満たす) 場合を考える。

→MNL形式が有効であれば、確定項のパラメータ θ^* の推定値が、全選択肢を列挙しなくても一部の選択肢集合から得られるという方法を確立したい。

条件設定(1)

真の選択確率を $P(i|C, z, \theta^*)$ とし、これはIIAを満たすと仮定する。

ただし、
C: 選択肢の完全な集合
z: 説明変数によるベクトル
 θ : パラメータのベクトル

すると、確率PはMNLモデルにより次のようになる。

$$i \in D \subseteq C \rightarrow P(i|C, z, \theta) = P(i|D, z, \theta) \sum_{j \in D} P(j|C, z, \theta) \quad (35)$$

条件設定(2)

確率 $\pi(D|i, z)$ に従って集合Cから部分集合Dが出されるとする。

ある選択肢iが観測されたとして、分布を以下のように設定。

(a) 選択肢iとは独立で、調整された部分集合D

(b) 選択肢iを含む、ランダムな部分集合D

(c-1) 選択肢iと、ランダムに選ばれた1つの選択肢という、
二要素から成る部分集合D

(c-2) Cを C_M に J_M 個の要素を持つ集合 $\{C_1, \dots, C_M\}$ に分割し、
分割集合 C_n からiを、残りの集合から無作為に一個ずつ
選択して形成されたD

もし選択肢jがDに存在する場合、jはDから選択される確率を持つ。

条件設定(2)

分割(c-1), (c-2)の確率を示す。

(c-1) 選択肢*i*と、ランダムに選ばれた1つの選択肢という、二要素から成る部分集合*D*。全選択肢数を*J*とする。

$$\pi(D|i, z) = \begin{cases} 1/(J - 1) & \text{if } D = \{i, j\} \text{ and } i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (36)$$

(c-2) *C*を*C_m*に*J_m*個の要素を持つ集合{*C₁*, ..., *C_M*}に分割し、分割集合*C_n*から*i*を、残りの集合から無作為に一個ずつ選択して形成された*D*

$$\pi(D|i, z) = \begin{cases} J_n / \prod_{m=1}^M J_n & \text{if } j \in D, M = \#(D), \forall m D \cap C_m \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (37)$$

特性

Positive Conditioning Property

$$j \in D \subseteq C \text{ and } \pi(D|i, z) > 0 \quad \rightarrow \quad \pi(D|j, z) > 0$$

Uniform Conditioning Property

$$i, j \in D \subseteq C \rightarrow \pi(D|i, z) = \pi(D|j, z)$$

先の四つのDは全てPositive Conditioning Propertyを満たす。
さらに(c-2)以外はより強い条件であるUniform Conditioning Propertyを満たす。

選択肢の中から一つ選ぶ行為が $n=1, \dots, N$ の N 回行われたとする。

この時選ばれた選択肢を i_n , それぞれのケースで部分集合を D_n とし、 C から D_n が抽出される確率を $\pi(D|i_n, z_n)$ とする。ただし、 i が D の中に無い選択は除外する。

この時、各回のMNLモデルは V_n を i_n の確定項として

$$P(i|C, z, \theta) = \exp[V_i(z, \theta)] / \sum_{j \in C} \exp[V_j(z, \theta)] \quad (38)$$

と表せる。

次に尤度を考える。

ここで考える尤度とは、パラメータ θ を置いた時に、集合 D から選択肢 i が選択される確率となる。

選択確率は式(18)より以下の通り。

$$P(i|C, z, \theta) = \sum_{m=1}^M P(i|D_m)P(D_m)$$

注目している集合D以外の集合にもiが入っている。よって
 選択肢iが選ばれてなおかつ集合Dが選ばれている

$$P(i|D, z, \theta) = \frac{\text{集合Dが選ばれている}}{\text{選択肢iが選ばれてなおかつ集合Dが選ばれている}}$$

$$= \frac{\text{選択肢jが選ばれなおかつ集合Dが選ばれている}}{\sum_{j \in D} \text{選択肢jが選ばれなおかつ集合Dが選ばれている}}$$

$$= \frac{P(i|C, z, \theta)\pi(D|i, z)}{\sum_{j \in D} P(j|C, z, \theta)\pi(D|j, z)}$$

$$= \frac{\{\exp[V_i(z, \theta)] / \sum_{k \in C} \exp[V_k(z, \theta)]\} \pi(D|i, z)}{\sum_{j \in D} \{\exp[V_j(z, \theta)] / \sum_{k \in C} \exp[V_k(z, \theta)]\} \pi(D|j, z)}$$

$$= \frac{\exp[V_i(z, \theta)] \pi(D|i, z)}{\sum_{j \in D} \exp[V_j(z, \theta)] \pi(D|j, z)}$$

$$= \frac{\exp[V_i(z, \theta)] \pi(D|i, z)}{\sum_{j \in D} \exp[V_j(z, \theta)] \pi(D|j, z)}$$

(続き)

$$\begin{aligned}
 P(i|D, z, \theta) &= \frac{\exp[V_i(z, \theta)]\pi(D|i, z)}{\sum_{j \in D} \exp[V_j(z, \theta)]\pi(D|j, z)} \\
 &= \frac{\exp[V_i(z, \theta) + \log \pi(D|i, z)]}{\sum_{j \in D} \exp[V_j(z, \theta) + \log \pi(D|j, z)]}
 \end{aligned}$$

これが尤度となる。

n=1~N回目の尤度の相乗平均を考えると

$$\left(\prod_{n=1}^N \frac{\exp[V_{i_n}(z_n, \theta) + \log \pi(D_n|i_n, z)]}{\sum_{j \in D} \exp[V_j(z_n, \theta) + \log \pi(D_n|j, z_n)]} \right)^{1/N}$$

この対数を取り、

$$L_N = (1/N) \sum_{n=1}^N \log \left\{ \frac{\exp[V_{i_n}(z_n, \theta) + \log \pi(D_n|i_n, z)]}{\sum_{j \in D} \exp[V_j(z_n, \theta) + \log \pi(D_n|j, z_n)]} \right\} \quad (39)$$

これが最大化するような θ を考え、 θ^* の推定値とする。

特に、Uniform Conditioning Propertyを満たす時、式(39)

$L_N = (1/N) \sum_{n=1}^N \log \left\{ \frac{\exp[V_{i_n}(z_n, \theta) + \log \pi(D_n | i_n, z)]}{\sum_{j \in D} \exp[V_j(z_n, \theta) + \log \pi(D_n | j, z_n)]} \right\}$ は次を満たす。

$$L_N = (1/N) \sum_{n=1}^N \log \left\{ \frac{\exp[V_{i_n}(z_n, \theta)]}{\sum_{j \in D} \exp[V_j(z_n, \theta)]} \right\} \quad (40)$$

以上より、選択過程がMNLモデルであれば選択肢を限定しても良く、(c-2)のようなメカニズムを使用する場合には式(39)のように修正する必要がある。

ただし(c-2)のような場合でも、選択メカニズムの効果が確定項のパラメータ内に吸収されていれば推定を行うことができる。

この場合、モデル内のクラス((c-2)における部分集合 C_m)変数の構成が選択メカニズムの効果に十分であることの検証が必要であり、このパラメータの推定値を解釈する際に注意が必要。

前節で、MNLモデルが使える場面では分析者により定義される各部分集合「クラス」から選択された「代表的な」選択肢を用いて推定を行うことができることを示した。これは、選択肢の集約の一つの方法である。

クラスの例：コミュニティ、住居タイプ

次に、観測できない属性間の依存性によりMNLモデルが使えない場合の集約方法を検討する。

設定

コミュニティ $c=1, \dots, C$ に N_c 個の住戸 $n=1, \dots, N_c$ があり、すべての住戸に共通の観測できないコミュニティ属性がある(IIA条件を満たしていない)ためにMNLモデルが使えない。

選択肢間の相互依存性を表すために、式(32)~(24)のNLモデルを用いる。

(復習)

$$\text{式(32)} \quad P_c = \exp[\alpha' y_c + (1 - \sigma)I_c] / \sum_{b=1}^C \exp[\alpha' y_b + (1 - \sigma)I_b]$$

$$\text{式(33)} \quad P_{n|c} = \exp[\beta' x_{cn} / (1 - \sigma)] / \exp(I_c)$$

$$\text{式(34)} \quad I_c = \log \sum_{m=1}^{N_c} \exp[\beta' x_{cm} / (1 - \sigma)]$$

ここで、 σ は選択肢間の類似の度合いであり、 $\sigma=0$ の場合MNLモデルとなる。

$\sigma=1$ の場合、以下のようなになる。

$$P_c = \exp[\alpha' y_c + \max \beta' x_{cn}] / \sum_{b=1}^C \exp[\alpha' y_b + \max \beta' x_{bn}] \quad (41)$$

$$\text{式(41)} P_c = \exp[\alpha' y_c + \max \beta' x_{cn}] / \sum_{b=1}^C \exp[\alpha' y_b + \max \beta' x_{bn}]$$

Quigley(1976)はピッツバーグのデータを用いて、式(41)のモデルで住宅需要を分析している。(ただしコミュニティタイプと住戸タイプが逆)

彼はパラメータ推定に先立ち $\max \beta' x_{cn}$ が計算できるように立地選択の構造を仮定しており、こうすることで**式(41)をMNLモデルとして扱うことができる。**

→このモデルでは相関を持つグループ内で代表を決めることができれば、MNLモデルに持ち込むことができる。

Lerman(1977)のワシントンD.C.のデータwp用いた分析では、以下の形でモデルを推定している。

$$P_c = \frac{\exp[\alpha' y_c + X_c^* + (1 - \sigma) \log N_c]}{\sum_{b=1}^C \exp[\alpha' y_b + X_b^* + (1 - \sigma) \log N_b]} \quad (42)$$

ただし、

- ・ c は国勢統計区を表す
- ・ X_c^* は統計区 c 内の住戸の効用 $\beta' x_{cn}$ の平均を表す
- ・ $\log N_c$ は、大きく住戸数の多い統計区は選択される確率が高くなるだろうという見立てで、補正のために入れられた大きさの尺度である。なぜ \log になるかは後述。

ゾーン内の住戸が比較的均質で、MNLモデルが適用できる場合、地区サイズ補正 $\log N_b$ の係数 $(1 - \sigma)$ は1となる。参考：式(39)

式(41) $P_c = \exp[\alpha' y_c + \max \beta' x_{cn}] / \sum_{b=1}^C \exp[\alpha' y_b + \max \beta' x_{bn}]$ は
式(42) $P_c = \frac{\exp[\alpha' y_c + X_c^* + (1-\sigma) \log N_c]}{\sum_{b=1}^C \exp[\alpha' y_b + X_b^* + (1-\sigma) \log N_b]}$ の極端なケースである。

Lermanの調査によると

また、MNLモデルの仮定(IIA条件)が満たされない場合、係数 $(1-\sigma)$ は1とならず、実際には $1-\sigma = 0.492 \pm 0.094$ としている。

→ σ は0とも1とも有意に異なる。

一方、式(34) $I_c = \log \sum_{m=1}^{N_c} \exp[\beta' x_{cn}/(1 - \sigma)]$ は次のように書き直すことができる。

$$I_c = [X_c^*/(1 - \sigma)] + \log N_c + \log \frac{1}{N_c} \sum_{m=1}^{N_c} \exp \left[\frac{\beta' x_{cn} - X_c^*}{1 - \sigma} \right] \quad (43)$$

式(42)はここが省略された形だった

統計区c内で住戸がすべて観測変数において一様であれば、全ての住戸nで $\beta' x_{cn} - X_c^* = 0$ となり、NLモデルは式(43)の最後の項が無くなり、Lermanモデル(式42)となる。

よって

- ・ 式(43)は確率的効用最大化とLermanモデルの整合性を確立
- ・ Lerman 「 $\log N_c$ の係数が独立の度合いを示す」を支持

また、 $X_c^* = \max \beta' x_{cn}$ とするとQuigleyのモデル(式(41))になる。

$$\text{式(43)} \quad I_c = [X_c^*/(1 - \sigma)] + \log N_c + \log \frac{1}{N_c} \sum_{m=1}^{N_c} \exp \left[\frac{\beta' x_{cn} - X_c^*}{1 - \sigma} \right]$$

X_c^* を $\beta' x_{cn}$ の平均、ただし $(\forall n \ \beta' x_{cn} = X_c^*)$ では無いとすると、指数関数の凸性から $\exp[(\beta' x_{cn} - X_c^*)/(1 - \sigma)]$ の平均は

$$\frac{1}{N_c} \sum_{m=1}^{N_c} \exp \left[\frac{\beta' x_{cn} - X_c^*}{1 - \sigma} \right] > 1 \quad (44)$$

→式(43)より $I_c > [X_c^*/(1 - \sigma)] + \log N_c$ となり、両辺の差は $\beta' x_{cn}$ の分散に依存する。

一つの極限を考える。

(条件)

統計区内の住戸数が多い。 $N_c = r_c N$ とおき、 r_c を固定で $N \rightarrow \infty$ $\beta' x_{cn}$ が平均 X_c^* 分散 ω_c^2 の独立した正規分布とする。

この時、

$$P_c \rightarrow \frac{\exp\left\{\left[\alpha' y_c + \beta' x_c^* + (1-\sigma) \log r_c + \frac{1}{2} \omega_c^2\right] / (1-\sigma)\right\}}{\sum_{b=1}^C \exp\left\{\left[\alpha' y_b + \beta' x_b^* + (1-\sigma) \log r_b + \frac{1}{2} \omega_b^2\right] / (1-\sigma)\right\}} \quad (45)$$

非集計データ x_{cn} が観測できないが、その分布が近似または推計でき、 ω が既知である場合、式(45)とMNLモデルの推定方法を用いて σ と β を推定することができる。

$$\text{式(45)} \quad P_c \rightarrow \frac{\exp\left\{\left[\alpha' y_c + \beta' x_c^* + (1-\sigma) \log r_c + \frac{1}{2} \omega_c^2\right] / (1-\sigma)\right\}}{\sum_{b=1}^C \exp\left\{\left[\alpha' y_b + \beta' x_b^* + (1-\sigma) \log r_b + \frac{1}{2} \omega_b^2\right] / (1-\sigma)\right\}}$$

非集計データ x_{cn} が観測できないが、その分布が近似または推計でき、 ω が既知である場合、式(45)とMNLモデルの推定方法を用いて σ と β を推定することができる。

しかし、 y_c に統計区固有のダミー変数がある場合
統計区固有の係数と r_c が不明である、

ただし、逆に統計区固有の係数の解釈の一つとして、統計区内の等価な選択肢数を部分的に示しているとも考えることもできる。

$$\text{式(45)} \quad P_c \rightarrow \frac{\exp\left\{\left[\alpha' y_c + \beta' x_c^* + (1-\sigma) \log r_c + \frac{1}{2} \omega_c^2\right] / (1-\sigma)\right\}}{\sum_{b=1}^C \exp\left\{\left[\alpha' y_b + \beta' x_b^* + (1-\sigma) \log r_b + \frac{1}{2} \omega_b^2\right] / (1-\sigma)\right\}}$$

・ ω_c は分からないが、 x_{cn} が共分散行列 Ω_c の多変量正規分布であり、 $\omega_c^2 = \beta' \Omega_c \beta$ であることが分かっている場合

→ この場合でも β, σ, α を直接推定することが可能

完全な最尤推定には、 β の非線形制約のために標準的なMNLモデルを修正する必要がある。

または、 $\beta' \Omega_c \beta$ の項を独立したパラメータとして書き出し、制約を無視することでも一定した推定値を得られる。