

sequential alimnet method によって統一的に評価することが可能となる。このような手法は、アクティビティパターンの生成過程を記述したりシミュレータのアクティビティパタン生成モデルに利用できる可能性がある。アクティビティを符号化して操作することで本来連続的な属性も扱いやすくなることが可能である。個々の属性の重みや各操作のコスト設定に関してはさまざまな設定が可能であり、より記述力の高い重み設定を工夫する余地もあるといえる。

## 3.2 行動モデル

### 3.2.1 離散選択モデル

(a) 効用と選択確率の定式化

#### 効用の定式化

離散選択モデルでは、選択肢ごとに効用をある関数により与えた上で、選択肢間の効用の比較によって選択結果が導出される。選択肢の効用は直接観測することはできないが、効用は連続変数であり、かつ説明変数の線形和であると仮定する。効用関数は次となる。

$$V_{in} = \beta_1 x_{1in} + \beta_2 x_{2in} + \cdots + \beta_K x_{Kin} \quad (3.120)$$

ここで、 $V_{in}$  は個人  $n$  の選択肢  $i$  に対する効用を示す。 $x_{kin}$  は個人  $n$  の選択肢  $i$  に対する  $k$  番目の説明変数、 $\beta_k$  は  $k$  番目の未知パラメータを示す。交通手段選択モデルの場合には、説明変数には所要時間・費用・乗り換え回数・交通手段までのアクセス時間等の交通手段別特性と、移動目的・性別・年齢等の個人ごとの移動・個人特性が用いられる。

説明変数 ( $x_{kin}$ ) の観測誤差や個人個人の重み付けパラメータ ( $\beta_k$ ) の差異が実際は存在しており、式 (3.120) の効用関数は測定可能な内容による近似に過ぎないと考えられる。そこで、真の効用関数は測定可能な効用と不可能な効用による和で表す。測定不可能な効用を確率変数を用いて表現し、真の効用関数は次となる。

$$\begin{aligned}
 U_{in} &= \beta_1 x_{1in} + \beta_2 x_{2in} + \cdots + \beta_K x_{Kin} + \epsilon_{in} \\
 &= V_{in} + \epsilon_{in}
 \end{aligned}
 \tag{3.121}$$

ここで、 $V_{in}$  は効用の確定部分 (測定可能な効用) を示し、 $\epsilon_{in}$  は効用の不確定部分 (効用の誤差項) を示す。  $U_{in}$  を確率効用 (random utility) という。行動者の真の効用を知ることができないために効用を確率変数を用いて表現している。効用の誤差項には次のような要因がある (北村ら (2002)[1] を参考に一部追加)。

- 確定部分に含まれる変数以外の要因 (抜け落ちた変数)
- 確定部分を線形和とした関数形の誤差
- 属性の重み  $\beta_k$  を個人間で均一とした誤差
- 説明変数の測定誤差
- 行動者の説明変数の認知誤差

#### 選択確率の定式化

個人  $n$  が選択肢  $i$  を選択するときは、選択肢  $i$  の効用  $U_{in}$  が他の選択肢の効用よりも大きいときであると考えられる。この選択確率  $P_n(i)$  は、式 (3.121) を用いて、次のように表現できる。

$$\begin{aligned}
 P_n(i) &= Pr[U_{in} \geq U_{jn}, \text{for } \forall j, j \neq i] \\
 &= Pr[V_{in} + \epsilon_{in} \geq V_{jn} + \epsilon_{jn}, \text{for } \forall j, j \neq i]
 \end{aligned}
 \tag{3.122}$$

これにより効用の誤差項の確率分布を特定化することで、選択確率  $P_n(i)$  を導出できる。

次に、選択可能な選択肢が 2 つの場合 (2 項選択問題) を例に、具体的に選択確率を導出する。まず、2 つの選択肢  $i, j$  について、式 (3.122) を変形する。

$$\begin{aligned}
P_n(i) &= Pr[U_{in} \geq U_{jn}] \\
&= Pr[V_{in} + \epsilon_{in} \geq V_{jn} + \epsilon_{jn}] \\
&= Pr[\epsilon_{jn} - \epsilon_{in} \geq V_{in} - V_{jn}] \quad (3.123) \\
&= Pr[\epsilon_n \geq V_{in} - V_{jn}] \\
&= F_\epsilon(V_{in} - V_{jn})
\end{aligned}$$

ここで、 $\epsilon_n \equiv \epsilon_{jn} - \epsilon_{in}$  とし、 $F_\epsilon$  は  $\epsilon_n$  の累積分布関数 (cumulative distribution function) である。選択確率の導出にあたっては、この誤差項の差である  $\epsilon_n$  に正規分布を仮定する場合とロジスティック分布を仮定する場合がある。正規分布を仮定する場合はプロビットモデル (probit model) と呼ばれる。前述の通り、効用の誤差項の要因には様々なものが含まれるため、この和の分布としては、中心極限定理により、正規分布を仮定するのが自然である。この場合の選択確率は次となる。

$$\begin{aligned}
P_n(i) &= \Phi_\epsilon(V_{in} - V_{jn}) \\
&= \int_{-\infty}^{V_{in} - V_{jn}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\epsilon}{\sigma}\right)^2\right] d\epsilon \quad (3.124) \\
&= \int_{-\infty}^{\frac{V_{in} - V_{jn}}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}z^2\right] dz \\
&= \Phi\left(\frac{V_{in} - V_{jn}}{\sigma}\right)
\end{aligned}$$

ここで、 $\sigma$  は  $\epsilon_n$  の標準偏差、 $\Phi$  は標準正規分布の累積分布関数を示す。このように、プロビットモデルでは、選択確率が積分形が残るオープンフォームとなり、計算負荷が高くなる。

対して、ロジスティック分布を仮定する場合は選択確率が積分形のないクローズドフォームとなる。これはロジットモデル (logit model) と呼ばれる。この場合の選択確率は次となる。

$$\begin{aligned}
P_n(i) &= \frac{1}{1 + \exp(-\mu(V_{in} - V_{jn}))} \\
&= \frac{\exp(\mu V_{in})}{\exp(\mu V_{in}) + \exp(\mu V_{jn})} \quad (3.125)
\end{aligned}$$

ここで、 $\mu$  はスケールパラメータであり、 $\epsilon_n$  のばらつきの程度を表し、標準偏差に反比例する値を取る。なお、誤差項の差である  $\epsilon_n$  にロジスティック分布を仮定する場合は、誤差項にはガンベル分布を仮定する。

また、2項選択問題において、プロビットモデルもロジットモデルも選択確率は図に表されるようなS字型となっている。

これ以降の (b)~(g) では誤差項にガンベル分布を仮定したクローズドフォームの離散選択モデル、(h)~?? では誤差項に正規分布を仮定したオープンフォームの離散選択モデルを説明する。

### (b) MNL モデル

MNL モデル (multinomial choice model(Luce(1959),McFadden(1973))) では、(a) で説明した2項選択問題を一般化し、選択枝数を3以上とした問題を扱うことができる。この問題は次のように考えられる。選択枝  $i$  の選択確率は選択枝  $i$  以外の中で最大の効用をもたらす選択枝よりも選択枝  $i$  の効用が大きい確率であり、定式化は次となる。

$$\begin{aligned} P_n(i) &= \Pr[U_{in} \geq U_{jn}, \text{ for } \forall j, j \neq i] \\ &= \Pr[U_{in} \geq \max_{\forall j, j \neq i} U_{jn}] \end{aligned} \quad (3.126)$$

多項ロジットモデルでは、それぞれの選択枝の誤差項に、独立で同一な (identically and independently distributed; IID) ガンベル分布 (第1種極値分布) を仮定している。ガンベル分布の分布形は次である。

<累積分布関数>

$$F(\epsilon) = \exp(-\exp(-\mu(\epsilon - \eta))) \quad (3.127)$$

<確率密度関数>

$$f(\epsilon) = \mu \exp(-\mu(\epsilon - \eta)) \exp(-\exp(-\mu(\epsilon - \eta))) \quad (3.128)$$

ここで、 $\mu$  はガンベル分布のスケールパラメータであり、 $\epsilon_n$  のばらつきの程度を表す。 $\eta$  は分布の位置 (最頻値) を表すロケーションパラメータである。平均値は  $\eta + \gamma/\mu$  (オイラー定数  $\gamma \simeq 0.577$ )、分散は  $\pi^2/6\mu^2$  である。

ここでガンベル分布の2つの性質を示す。性質(1)は次である。 $\epsilon_1$ と $\epsilon_2$ がパラメータ $(\eta_1, \mu), (\eta_2, \mu)$ に従うとき、 $\epsilon = \epsilon_1 - \epsilon_2$ は次のロジスティック分布に従う。

$$F(\epsilon) = \frac{1}{1 + \exp(\mu(\eta_2 - \eta_1 - \epsilon))} \quad (3.129)$$

性質(2)は次である。 $\epsilon_1, \epsilon_2, \dots, \epsilon_i, \dots, \epsilon_I$ がそれぞれパラメータ $(\epsilon_i, \mu)$ を持つ互いに独立なガンベル分布に従うとすると、 $\epsilon_1, \dots, \epsilon_I$ の最大値 $\max(\epsilon_1, \dots, \epsilon_I)$ もガンベル分布に従い、そのパラメータは次となる。

$$\left( \frac{1}{\mu} \ln \sum_{i=1}^I \exp(\mu\eta_i), \mu \right) \quad (3.130)$$

これらの性質を用いて多項ロジットモデルの選択確率の導出を行う。

まず、 $\epsilon_1, \epsilon_2, \dots, \epsilon_I$ がそれぞれパラメータ $(0, \mu)$ を持つ互いに独立なガンベル分布に従うとすると、式(3.126)の $\max_{j, j \neq i} U_{jn} \equiv U_n^*$ は性質(2)より次のパラメータに従うガンベル分布となる。

$$\left( \frac{1}{\mu} \ln \sum_{j \neq i} \exp(\mu V_{jn}), \mu \right) \quad (3.131)$$

ここで、 $U_n^* = V_n^* + \epsilon_n^*$ とし、 $V_n^* = \frac{1}{\mu} \ln \sum_{j \neq i} \exp(\mu V_{jn})$ とおく。これにより、 $\epsilon_n^*$ はパラメータ $(0, \mu)$ を持つガンベル分布に従う。

これを式(3.126)に代入し、性質(1)を用いると、次のように式変形できる。

$$\begin{aligned} P_n(i) &= Pr[V_{in} + \epsilon_{in} \geq V_n^* + \epsilon_n^*] \\ &= Pr[\epsilon_n^* - \epsilon_{in} \geq V_{in} - V_n^*] \\ &= \frac{1}{1 + \exp(\mu(V_n^* - V_{in}))} \\ &= \frac{\exp(\mu V_{in})}{\exp(\mu V_{in}) + \exp(\mu V_n^*)} \\ &= \frac{\exp(\mu V_{in})}{\exp(\mu V_{in}) + \exp\left(\ln \sum_{j \neq i} \exp(\mu V_{jn})\right)} \\ &= \frac{\exp(\mu V_{in})}{\sum_j \exp(\mu V_{jn})} \end{aligned} \quad (3.132)$$

これが多項ロジットモデルの選択確率となる。また、通常はスケールパラメータ  $\mu$  の値を 1 とおく。

次に、ロジットモデルの IIA 特性 (independence from irrelevant alternatives) について説明する。IIA とは無関係な選択肢から選択確率が独立であることである。例えば、2 つの選択肢  $i, j$  の選択確率の比は、 $P_{in}/P_{jn} = \exp(V_{in} - V_{jn})$  と表すことができる。この比率は選択肢  $i, j$  の効用の確定部分にのみ影響を受けており、 $i, j$  以外の選択肢からは影響を受けていない。これが IIA 特性である。

IIA 特性の長所としては、選択肢集合に含まれる全ての選択肢ではなく、その部分集合を用いてモデルを推定してもパラメータ推定値にバイアスが生じないことである。短所としては、類似した選択肢が存在し、選択肢間の誤差項が独立であるという仮定が誤っていた場合に類似選択肢の選択確率が過大評価されてしまうというものがある。この短所の説明にあたり、よく取り上げられる例が赤バス—青バス問題である。効用の確定部分の和が全く車と赤バスで同じだった場合に、その選択確率はどちらも  $1/2$  である。そこに、赤バスと全く同じ効用の確定部分をもつ青バスが導入された場合に、IIA 特性の下では選択確率は車  $1/3$ 、赤バス  $1/3$ 、青バス  $1/3$  となる。青バスが導入されてもバスの効用は変化しないため、選択確率が車  $1/2$ 、赤バス  $1/4$ 、青バス  $1/4$  となるのが直感的な答えである。このように、類似した選択肢が存在する場合に、それらの選択確率が過大評価されるのは選択肢間の誤差項が独立であると仮定しているためである。選択肢が類似している場合には、その誤差項の間にも相関があると考えられ、選択肢間の誤差項の相関を考慮したモデルが必要となる。多項ロジットモデルの IIA 特性を緩和したモデルの説明を (c)~(g) では行う。また、選択肢間の誤差項の相関の考慮が必要な例としては、交通手段選択問題におけるバスと鉄道の公共交通という相関、経路選択問題における同じリンクを経路に含んでいることによる相関等がある。

#### (c) NL モデル

IIA 特性を緩和したモデルとして、まず、ネスティッドロジットモデル (Nested Logit model: NL model (Ben-Akiva(1973))) を説明する。ここでは、目的エリアと交通手段の組合せの選択問題を例に説明する (図 (c))。上位

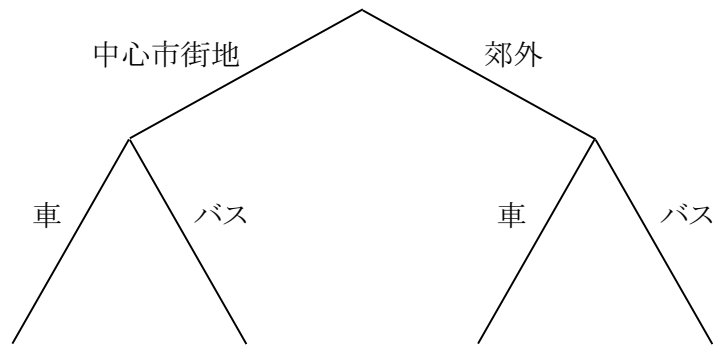


図 3.54 目的エリア・交通手段選択モデルのツリー図 (NL モデル型)

ネストである目的エリアの選択肢  $d$  は {中心市街地, 郊外}, 下位ネストである交通手段の選択肢  $i$  は {車, バス} とする. 選択肢は 1:( $d$  = 中心市街地,  $i$  = 車), 2:(中心市街地, バス), 3:(郊外, 車), 4:(郊外, バス) の4つとなる. 目的エリアごとの観測できない選択要因 (効用の誤差項) があるため, 選択肢 1 と 2, 選択肢 3 と 4 の下位ネストの中で誤差項の相関が生まれる. これを定式化すると次となる. なお, ここでの説明では個人を表す  $n$  の添え字は省略する.

$$U_{di} = V_d + V_i + V_{di} + \epsilon_d + \epsilon_{di} \quad (3.133)$$

ここで,  $U_{di}$  は選択肢  $di$  の組合せの効用,  $V_d, V_i$  は選択肢  $d, i$  のそれぞれに特有の効用の確定部分,  $V_{di}$  は  $d, i$  の組合せで決まる効用の確定部分とする.  $\epsilon_d$  は選択肢  $d$  に特有の効用の誤差項 ( $\max U_{di}$  がスケールパラメータ  $\mu^d$  を持つガンベル分布になるような分布に従うと仮定) とする.,  $\epsilon_{di}$  は  $d, i$  の組合せで決まる効用の誤差項 (スケールパラメータ  $\mu$  をもつ互いに独立なガンベル分布に従うと仮定) とする. 下位ネストの選択肢は誤差項として  $\epsilon_d$  を共通に持つこととなる.

次に, 選択肢  $di$  の選択確率  $P(d, i)$  を導出する. 選択確率は, 条件つき確率  $P(i|d)$  と周辺確率  $P(d)$  の積によって表す.

$$P(d, i) = P(i|d)P(d) \quad (3.134)$$

また, 周辺確率は次となる.

$$\begin{aligned}
P(d) &= Pr \left[ \max_i U_{di} \geq \max_i U_{d'i}, d' \neq d \right] \\
&= Pr \left[ V_d + \epsilon_d + \max_i (V_i + V_{di} + \epsilon_{di}) \right. \\
&\quad \left. \geq V_{d'} + \epsilon_{d'} + \max_i (V_i + V_{d'i} + \epsilon_{d'i}), d' \neq d \right]
\end{aligned} \tag{3.135}$$

$\epsilon_{di}$  の分布の仮定より,  $\max_i (V_i + V_{di} + \epsilon_{di})$  もスケールパラメータ  $\mu$  を持つガンベル分布に従う. また, このガンベル分布のロケーションパラメータ  $V'_d$  は, 前述の性質より,  $V'_d \equiv \frac{1}{\mu} \ln \sum_i \exp(\mu(V_i + V_{di}))$  となる.  $V'_d$  はログサム変数と呼ばれる. ここで, 式(3.135)は次となる.

$$P(d) = Pr \left[ V_d + V'_d + \epsilon_d + \epsilon'_d \geq V_{d'} + V'_{d'} + \epsilon_{d'} + \epsilon'_{d'}, d' \neq d \right] \tag{3.136}$$

なお,  $\epsilon'_d$  は  $\epsilon'_d \equiv \max_i (V_i + V_{di} + \epsilon_{di}) - V'_d$  とする. 式(3.136)は, 効用の確定部分として  $V_d + V'_d$ , 誤差項として  $\epsilon_d + \epsilon_{d'}$  を与えた離散選択モデルの構造となっている. ここで,  $\epsilon_d$  の仮定より, 周辺確率は次の式で与えられる.

$$P(d) = \frac{\exp(\mu^d (V_d + V'_d))}{\sum_{d'} \exp(\mu^d (V_{d'} + V'_{d'}))} \tag{3.137}$$

次に, 条件つき確率  $P(m|d)$  (目的エリア  $d$  が決まった場合の交通手段  $i$  の選択確率) は, 次となる.

$$\begin{aligned}
P(m|d) &= Pr \left[ U_{di} \geq U_{d'i'}, i' \neq i | d \right] \\
&= Pr \left[ V_i + V_{di} + \epsilon_{di} \geq V_{i'} + V_{d'i'} + \epsilon_{d'i'}, i' \neq i | d \right]
\end{aligned} \tag{3.138}$$

下位ネストの中で目的エリア  $d$  は共通しており,  $V_d$  や  $\epsilon_d$  は選択に影響を与えない.  $\epsilon_{di}$  は IID ガンベル分布を仮定しており, 条件つき確率  $P(m|d)$  は次となる.

$$P(i|d) = \frac{\exp(\mu(V_i + V_{di}))}{\sum_{i'} \exp(\mu(V_{i'} + V_{d'i'}))} \tag{3.139}$$

式(3.134), 式(3.137), 式(3.139)より, 各選択枝の選択確率  $P(d, i)$  は次となる.



$$\begin{aligned}
P(d, i) &= P(i|d)P(d) \\
&= \frac{\exp(\mu(V_i + V'_{di}))}{\sum_{i'} \exp(\mu(V_{i'} + V'_{di'}))} \frac{\exp(\mu^d(V_d + V'_d))}{\sum_{d'} \exp(\mu^d(V_{d'} + V'_{d'}))} \quad (3.140)
\end{aligned}$$

式 (3.140) において、2つのスケールパラメータ  $\mu^d$  と  $\mu$  を同時に定めることはできない。そこで、条件つき確率のスケールパラメータ  $\mu$  を 1 とし、周辺確率のスケールパラメータ  $\mu^d$  を求めることが多い。ガンベル分布のスケールパラメータは分布のばらつきの程度を表しており、標準偏差に反比例する値を取るため、 $\mu^d/\mu$  は次式となる。

$$\frac{\mu^d}{\mu} = \frac{\sqrt{\frac{1}{\text{Var}(\epsilon_d + \epsilon'_d)}}}{\sqrt{\frac{1}{\text{Var}(\epsilon_{di})}}} = \sqrt{\frac{\text{Var}(\epsilon_{di})}{\text{Var}(\epsilon_d) + \text{Var}(\epsilon_{di})}} \leq 1 \quad (3.141)$$

$\epsilon'_d$  の定義よりスケールパラメータは  $\mu$  であり、これは  $\epsilon_{di}$  と同じことを用いている。これより、図 (c) の上位下位のネスト構造の仮定 (誤差項の相関構造の仮定) が正しいのであれば、 $\mu^d/\mu$  は 1 以下でなければならない。

#### (d) GEV モデル

ここで、クローズドフォームの離散選択モデルの定式化について、一般的なモデルである GEV モデル (General Extreme Value model, McFadden(1978)) を説明する。GEV ファミリーのモデル (本項の (b)~(g)) は GEV 理論から導出でき、GEV ファミリーと呼ばれる。

GEV モデルにおいて、選択肢集合  $C = (1, \dots, i, \dots, n)$  の中から選択肢  $i$  を選ぶ選択確率  $P(i|C)$  は次式で表される。

$$\begin{aligned}
P(i|C) &= \frac{y_i \frac{\partial G(y_1, y_2, \dots, y_n)}{\partial y_i}}{\mu G(y_1, y_2, \dots, y_n)} \quad (3.142) \\
y_i &= \exp(V_i), (U_i = V_i + \epsilon_i, i = 1, 2, \dots, n)
\end{aligned}$$

ここで、 $n$  は選択肢数、 $G$  は  $\mu$ -GEV 関数である。 $\mu$ -GEV 関数は次の性質を持つ微分可能な関数である。

1.  $G(\mathbf{y}) \geq 0$  for all  $\mathbf{y} \in \mathbb{R}_+^J$
2.  $G(\mathbf{y})$  は  $\mu$  次の同次関数である。  $G(\lambda \mathbf{y}) = \lambda^\mu G(\mathbf{y}), \lambda > 0$
3.  $\lim_{y_i \rightarrow \infty} G(y_1, \dots, y_i, \dots, y_n) = +\infty$ , for each  $i = 1, \dots, n$

4.  $G(\mathbf{y})$  の混合偏導関数が存在し、かつ混合偏導関数は連続である。  $k$  階偏導関数  $D_\kappa(\mathbf{y})$  は次が成立する。

$$(-1)^k D_\kappa(\mathbf{y}) \leq 0, \forall \mathbf{y} \in \mathbb{R}_+^J \quad (3.143)$$

$$\kappa = (i_1, \dots, i_k), D_\kappa(\mathbf{y}) = \frac{\partial^k G}{\partial y_{i_1} \dots \partial y_{i_k}}(\mathbf{y})$$

これらの性質から式 (3.142) が選択確率を示すことは次のように証明される。

**証明**

関数  $F$  を次のように定義する。

$$F(\epsilon_1, \dots, \epsilon_j, \dots, \epsilon_J) = \exp(-G(e^{-\epsilon_1}, \dots, e^{-\epsilon_j}, \dots, e^{-\epsilon_J})) \quad (3.144)$$

関数  $F$  が累積分布関数であり、ガンベル分布の多変量極値分布であることを示す。性質 3 より、関数  $F$  は次の性質をもつ。

$$\lim_{\epsilon_j \rightarrow -\infty} G = +\infty \text{ then } \lim_{\epsilon_j \rightarrow -\infty} F = 0 \quad (3.145)$$

$$\lim_{\{\epsilon_j\} \rightarrow +\infty} G = 0 \text{ then } \lim_{\{\epsilon_j\} \rightarrow +\infty} F = 1$$

負の極限が 0、正の極限が 1 に一致しており、累積分布関数の性質を満たす。次に関数  $Q_k$  を次のように定義する。

$$Q_1 = G_1 = \frac{\partial G(y_1, y_2, \dots, y_n)}{\partial y_1}$$

$$Q_k = Q_{k-1} G_k - \partial Q_{k-1} / \partial y_k \quad (3.146)$$

ここで、帰納的に  $Q_k$  が非負であることを示す。  $Q_{k-1}$  を非負とすると、  $G$  の 1 階微分である  $G_k$  は性質 4 より非負であるため、  $Q_{k-1} G_k$  は非負となる。  $\partial Q_{k-1} / \partial y_k$  を非正とする。このとき、  $\partial Q_k / \partial y_{k+1}$  は次式となる。

$$\frac{\partial Q_k}{\partial y_{k+1}} = \frac{\partial Q_{k-1}}{\partial y_{k+1}} G_k + Q_{k-1} \frac{\partial G_k}{\partial y_{k+1}} - \frac{\partial^2 Q_{k-1}}{\partial y_k \partial y_{k+1}} \quad (3.147)$$

偏微分を行うたびに正負が入れ替わる性質 4 より、  $\partial Q_k / \partial y_{k+1}$  は非正となる。以上より、  $Q_k$  は非負となる。この結果を用いて、  $F$  の偏微分は非負であることを帰納的に示す。1 階偏微分は次となる。

$$\begin{aligned}
\frac{\partial F}{\partial \epsilon_1} &= \exp(-G(e^{-\epsilon_1}, \dots, e^{-\epsilon_J})) \cdot \frac{\partial}{\partial \epsilon_1} (-G(e^{-\epsilon_1}, \dots, e^{-\epsilon_J})) \\
&= \mathbf{F} \cdot (-G_1) \cdot (-e^{-\epsilon_1}) \\
&= e^{-\epsilon_1} Q_1 \mathbf{F} \geq 0
\end{aligned} \tag{3.148}$$

( $k - 1$ ) 階偏微分を次とする.

$$\frac{\partial^{k-1} F}{\partial \epsilon_1 \cdots \partial \epsilon_{k-1}} = e^{-\epsilon_1} \cdots e^{-\epsilon_{k-1}} Q_{k-1} \mathbf{F}$$

$k$  階偏微分は次式となる.

$$\begin{aligned}
\frac{\partial^k F}{\partial \epsilon_1 \cdots \partial \epsilon_k} &= \frac{\partial}{\partial \epsilon_k} (e^{-\epsilon_1} \cdots e^{-\epsilon_{k-1}} Q_{k-1} \mathbf{F}) \\
&= (e^{-\epsilon_1} \cdots e^{-\epsilon_{k-1}}) \left( \frac{\partial Q_{k-1}}{\partial \epsilon_k} \mathbf{F} + Q_{k-1} \frac{\partial \mathbf{F}}{\partial \epsilon_k} \right) \\
&= e^{-\epsilon_1} \cdots e^{-\epsilon_k} Q_k \mathbf{F} \geq 0
\end{aligned} \tag{3.149}$$

ここでは、式 (3.146) と式 (3.148) を用いている。よって、 $\mathbf{F}$  の微分は単調増加であり、かつ連続である。これらの性質から  $\mathbf{F}$  は累積分布関数であるといえる。次に、例えば、 $i$  以外の  $j$  について  $\epsilon_j = +\infty$  を与えたとき、 $\mathbf{F} = \exp(-a_i \exp(-\epsilon_i))$  となる (なお、 $a_i = G(0, \dots, 0, i = 1, 0, \dots, 0)$ )。これは、スケールパラメータ  $\mu$  を 1、ロケーションパラメータ  $\eta$  を 0 としたときのガンベル分布と一致する。以上より、 $\mathbf{F}$  は多変量極値分布である。

選択枝  $j$  の効用を  $U_j = V_j + \epsilon_j$  としたときに、選択枝  $i$  の選択確率  $P(i)$  は次となる。

$$\begin{aligned}
P(i) &= \int_{-\infty}^{+\infty} \frac{\partial F(\dots, -\epsilon - V_i + V_{i-1}, -\epsilon, -\epsilon - V_i + V_{i+1}, \dots)}{\partial \epsilon_i} d\epsilon \\
&= \int_{-\infty}^{+\infty} e^{-\epsilon} G_i(\dots, e^{-\epsilon - V_i + V_{i-1}}, e^{-\epsilon}, e^{-\epsilon - V_i + V_{i+1}}, \dots) \cdot \\
&\quad \exp\left(-G(\dots, e^{-\epsilon - V_i + V_{i-1}}, e^{-\epsilon}, e^{-\epsilon - V_i + V_{i+1}}, \dots)\right) d\epsilon \\
&= \int_{-\infty}^{+\infty} e^{-\epsilon} G_i(\dots, e^{V_{i-1}}, e^{V_i}, e^{V_{i+1}}, \dots) \cdot \\
&\quad \exp\left(-e^{-\epsilon} G(\dots, e^{V_{i-1}}, e^{V_i}, e^{V_{i+1}}, \dots)\right) d\epsilon \\
&= e^{V_i} \frac{G_i(\dots, e^{V_{i-1}}, e^{V_i}, e^{V_{i+1}}, \dots)}{G(\dots, e^{V_{i-1}}, e^{V_i}, e^{V_{i+1}}, \dots)} \quad (3.150)
\end{aligned}$$

式展開では、 $G$  が同次関数であるという性質 2 を用いている ( $\mu = 1, \lambda = e^{-\epsilon} e^{-V_i}$ )。以上より、式 (3.150) は式 (3.142) に一致し、式 (3.142) が選択確率を示すことが証明された。■

$\mu$ -GEV 関数  $G(\mathbf{y})$  が次式で表されるとき、選択確率  $P(i|C)$  は MNL モデル、NL モデルと一致する。

$$\begin{aligned}
MNL : G(\mathbf{y}) &= \sum_{i=1}^J y_i^\mu \\
NL : G(\mathbf{y}) &= \sum_{d=1}^D \left( \sum_{i=1}^{J_i} y_i^{\mu_d} \right)^{\frac{\mu}{d}}
\end{aligned}$$

#### (e) CNL モデル

IIA 特性を緩和したモデルとして、次に、クロスネステッドロジットモデル (Cross Nested Logit model: CNL model) を説明する。CNL モデルは、選択肢のネストへの帰属度を設定できる。また、複数のネストに帰属することも可能である。この二つの点で NL モデルとは異なり、選択肢とネストの関係を帰属度により構造化できることに特徴がある。

具体的に、交通手段の選択問題を例に説明する (図 3.55)。交通手段の選択肢  $i$  は {私有車, バス, 鉄道} とする。上位ネスト  $m$  を {自動車, 公共交通} とする。バスは自動車であるために私有車と誤差項の相関を持つ。また、バスは公共交通でもあるために鉄道とも誤差項の相関を持つ。そこで、バスは

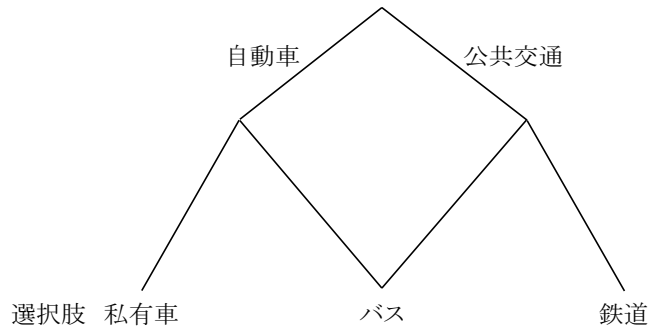


図 3.55 交通手段選択モデルのツリー図 (CNL モデル型)

自動車・公共交通の両方のネストに帰属すると考えられる。バスの自動車・公共交通への帰属度はアロケーションパラメータ  $\alpha_{im}$  により表す。バスの自動車への帰属度と公共交通への帰属度の和は1となる。

次に、 $\mu$ -GEV 関数は次のように定式化される。

$$G(y_1, \dots, y_n) = \sum_{m=1}^M \left( \sum_{j \in C} (\alpha_{jm}^{1/\mu} y_j)^{\mu_m} \right)^{\frac{\mu}{\mu_m}} \quad (3.151)$$

また、選択枝集合  $C$  における選択枝  $i$  の選択確率  $P(i|C)$  は次となる。ここで、 $\alpha_{im}$  はアロケーションパラメータであり、選択枝  $i$  のネスト  $m$  への帰属度を表す。また、 $M$  はネスト数である。 $\mu_m$  はネスト  $m$  のスケールパラメータ、 $\mu$  は選択枝  $i$  のスケールパラメータであり、 $0 < \mu \leq \mu_m$  となる ( $\mu_m = 1$  とおくことが多い)。

$$P(i|C) = \sum_{m=1}^M \frac{\left( \sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j} \right)^{\frac{\mu}{\mu_m}}}{\sum_{m'=1}^M \left( \sum_{j \in C} \alpha_{jm'}^{\mu_{m'}/\mu} e^{\mu_{m'} V_j} \right)^{\frac{\mu}{\mu_{m'}}}} \cdot \frac{\alpha_{im}^{\mu_m/\mu} e^{\mu_m V_i}}{\sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}} \quad (3.152)$$

また、ネスト  $m$  の選択確率  $P_m$ 、ネスト  $m$  が選択された上での選択枝  $i$  の条件つき選択確率  $P_{i|m}$  は次となる。

$$P(i|C) = \sum_{m=1}^M P_m P_{i|m} \quad (3.153)$$

$$P_m = \frac{\left( \sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j} \right)^{\frac{\mu}{\mu_m}}}{\sum_{m'=1}^M \left( \sum_{j \in C} \alpha_{jm'}^{\mu_{m'}/\mu} e^{\mu_{m'} V_j} \right)} \quad (3.154)$$

$$P_{i|m} = \frac{\alpha_{im}^{\mu_m/\mu} e^{\mu_m V_i}}{\sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}} \quad (3.155)$$

アロケーションパラメータ  $\alpha_{im}$  は次の条件を満たす。

$$0 \leq \alpha_{im} \leq 1, \sum_m \alpha_{im} = 1, \forall i \quad (3.156)$$

$\alpha_{im}$  は推定可能なパラメータであり、構造化を行うことが可能である。一方で、予め  $\alpha_{im}$  を与えた上で、効用  $V_i$  を推定するケースも多い。

(f) GNL モデル

GNL モデルは、CNL モデルや PCL モデル (Chu, 1989[2]) の一般化モデルとして開発されたモデルである。

まず、定式化の内容を示す。  $\mu$ -GEV 関数は次となる。

$$G(y_1, y_2, \dots, y_n) = \sum_m \left( \sum_{i' \in N_m} (\alpha_{i'm} y_{i'})^{1/\mu_m} \right)^{\mu_m} \quad (3.157)$$

ここで、 $N_m$  はネスト  $m$  下にある選択肢の集合である。  $\mu_m$  は、ネスト  $m$  のスケールパラメータであり、  $0 < \mu_m \leq 1$  である。  $\alpha_{im}$  は選択肢  $i$  のネスト  $m$  への帰属度を示すアロケーションパラメータである。なお、  $\alpha_{im}$  が満たすべき条件は次となる。

$$\alpha_{im} \geq 0, \sum_m \alpha_{im} = 1$$

ここで、選択肢  $i$  の選択確率  $P_i$  は次となる。

$$P_i = \frac{\sum_m \left( (\alpha_{im} e^{V_i})^{1/\mu_m} \left( \sum_{i' \in N_m} (\alpha_{i'm} e^{V_{i'}})^{1/\mu_m} \right)^{\mu_m - 1} \right)}{\sum_m \left( \sum_{i' \in N_m} (\alpha_{i'm} e^{V_{i'}})^{1/\mu_m} \right)^{\mu_m}} \quad (3.158)$$

また、選択確率は条件つき確率を用いて、次のように分解できる。

$$P_i = \sum_m P_{i|m} P_m \quad (3.159)$$

$$P_m = \frac{\left(\sum_{i' \in N_m} (\alpha_{i'm} e^{V_{i'}})^{1/\mu_m}\right)^{\mu_m}}{\sum_m \left(\sum_{i' \in N_m} (\alpha_{i'm} e^{V_{i'}})^{1/\mu_m}\right)^{\mu_m}} \quad (3.160)$$

$$P_{i|m} = \frac{(\alpha_{im} e^{V_i})^{1/\mu_m}}{\sum_{i' \in N_m} (\alpha_{im} e^{V_i})^{1/\mu_m}} \quad (3.161)$$

直接弾力性と交差弾力性の説明. Discrete Choice Analysis. The MIT Press, Cambridge, MA

$$\frac{\sum_m P_m P_{i|m} \left( (1 - P_i) + \left(\frac{1}{\mu_m} - 1\right) (1 - P_{i|m}) \right)}{P_i} \beta X_i \quad (3.162)$$

$$- \left( P_i + \frac{\sum_m \left(\frac{1}{\mu_m}\right) P_m P_{i|m} P_{i'|m}}{P_{i'}} \right) \beta X_i \quad (3.163)$$

(g) network GEV モデル

Daly and Bierlaire(2006) では、network GEV モデルとして、選択肢間の相関構造をネットワーク構造で表現することで、GEV モデルを開発できることが示された [3]。GEV モデルを開発するには、 $\mu$ -GEV 関数のを満たすことの証明が必要であったが、network GEV モデルを用いることで、この証明が簡易化される。

まず、GEV-network の説明を行う。GEV-network の任意のノードにおいて、GEV モデルが適用すれば、効用最大化理論と一致する。そのため、モデル内の潜在的な誤差項の相関を表現する GEV-network をデザインすることで GEV モデルを作ることができる。N をノードの集合、A をリンクの集合とし、G(N, A) を有限非空有向グラフとする。G(N, A) はネットワークであり、回路を含まないとする。また、リンク (i, j) は非負の associated parameter  $\alpha_{ij}$  をもつ。このネットワークを GEV-network と呼ぶ。例えば、

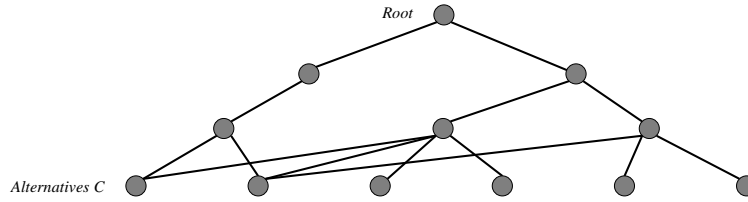


図 3.56 GEV-network の例

図 (g) は GEV-network である。ここでは最下層は選択肢集合  $C$  である。また、図 (g) は最上層のノードを 1 つとして、シングルルートネットワークである。

GEV-network 内のノード  $v_i \in C$  の  $G$  関数  $G^i$  は次となる。

$$G^i(y_i) = \sum_{v_j \in S(v_i)} \alpha_{ij} G^j(y)^{\frac{\mu_i}{\mu_j}} \quad (3.164)$$

関数  $G^i$  は、GEV 関数となる。これは、GEV 関数の累乗の線型結合は GEV 関数であるという性質に基づいている。この詳細の照明は、Daly and Bierlaire(2006)[3] を参照いただきたい。なお、 $S(v_i)$  はノード  $v_i$  の下層ネットワークに含まれるノードの集合を示す。以上のように、ネットワーク構造を設定することで、マルチクラスのネストによる GEV モデルを構築することができる。

(h) MNP モデル

次に、誤差項に正規分布を仮定する多項プロビットモデル (multinomial probit; MNP) を説明する [4][5]。本節の冒頭で説明した通り、効用の誤差項の要因には様々なものが含まれるため、中心極限定理より、誤差項には正規分布を仮定するのが自然である。そのため、MNP モデルのほうが、GEV モデルに比べて、誤差項の仮定が緩いモデルである。

個人  $n$  の選択肢  $i$  の MNP モデルの効用項と誤差項は次となる。

$$U_{in} = V_{in} + \epsilon_{in}, \quad i = 1, \dots, I \quad (3.165)$$

$$\epsilon_n = (\epsilon_{1n}, \epsilon_{2n}, \dots, \epsilon_{In}) \quad (3.166)$$

誤差項が平均 0、分散共分散行列が  $\Omega$  である多変量正規分布を与えたモデルである。選択肢間の異分散性を自由に表現することができる。選択肢  $i$  の選択確率は次となる。



$$P(i) = \int_{\epsilon_1=-\infty}^{\epsilon_i+V_i-\epsilon_1} \cdots \int_{\epsilon_i=-\infty}^{\infty} \cdots \int_{\epsilon_J=-\infty}^{\epsilon_i+V_i-\epsilon_J} \phi(\epsilon) d\epsilon_J \cdots \quad (3.167)$$

$$\phi(\epsilon) = \frac{1}{2\pi^{\frac{I-1}{2}} |\sigma|^{1/2}} \exp\left(-\frac{1}{2} \epsilon \sigma^{-1} \epsilon'\right) \quad (3.168)$$

表現力は高いものの、選択枝数-1の多重積分が必要であり、パラメータ推定が煩雑である。MNPの推定方法は乱数近似によるシミュレーション法やベイズ推定といった手法の導入により30程度の選択枝数でも計算可能となっている。推定法については、??推定法の節で詳細は説明する。また、MACML(Maximum Approximate Composite Marginal Likelihood)[6][7]は、多変量累積標準正規分布と合成周辺尤度を用いることでopen-formの離散選択モデルを簡単かつ高速で推定する手法である。この方法によれば、10以上の選択枝の場合も計算可能とされている。

(i) MXL(MMNL)モデル

Mixed Logitモデルはプロビットモデルに代表される正規分布を誤差項に仮定したモデルと、MNLモデルに代表されるガンベル分布を誤差項に仮定したモデルを組み合わせたモデルである。MNLモデルを誤差項操作により、変数の系列相関や分散を表現できるように拡張したモデルであるといえる。正規分布を確率項に用いたモデルでは、選択確率式がクローズドフォームになり解析的に求めることができない問題があったが、計算機の性能の向上により、シミュレーション計算による確率分布からのサンプリングにより、計算が実用的に可能となった。

Mixed Logitモデルの特徴は、母集団に対してパラメータを、確率分布により期待値と分散を用いて扱うことにより、母集団内でのばらつきを表現できることである。また、従前のモデルでは扱うことのできなかった誤差項の相関や異分散性を自由に表現できることが利点として挙げられる。一方で、多重積分を求めるためのシミュレーション近似計算は計算負荷が大きく、推定に要するコストが大きいという点が指摘される。

モデルの概要を示す。Mixed Logitモデルでは個人 $n$ が選択枝 $i$ を選ぶ効用関数は以下のように表される。

$$U_{in} = V_{in} + \eta_{in} + \epsilon_{in} \quad (3.169)$$

$$V_{in} = \beta_i + X_{in} \quad (3.170)$$

ただし、 $V_{in}$  はパラメータ  $\beta$ ，説明変数ベクトル  $\mathbf{X}$  で表される効用の確定項， $\eta$  は共分散行列  $\Omega$  に従う確率分布に従う誤差項， $\epsilon$  は IID ガンベル分布に従う誤差項である。パラメータはパラメータ値が確率的に変動するランダム係数として表され、ランダム係数がサンプル間の異質性を考慮した部分になる。Mixed Logit モデルでは、誤差項は正規分布に従う誤差項とガンベル分布に従う誤差項 2 つに分解される。そのうち前者は、誤差項の構造化やパラメータごとの固有の誤差項への分割が可能である。

個人  $n$  が選択肢  $i$  を選択する選択確率  $P_{in}$  は、効用関数と確率分布に従う  $\eta_{in}$  の値が与えられたとき、以下のようになる。

$$P_{in}(\eta_{in}) = \frac{\exp(V_{in} + \eta_{in})}{\sum_j \exp(V_{jn} + \eta_{jn})} \quad (3.171)$$

$\eta_{in}$  は確率分布であるため、分散共分散行列  $\Omega$  に従う確率密度関数  $f(\eta_{in}|\Omega)$  を用いて選択確率を求める。

$$P_{in} = \int \frac{\exp(V_{in} + \eta_{in})}{\sum_j \exp(V_{jn} + \eta_{jn})} f(\eta_{in}|\Omega) d\eta_{in} \quad (3.172)$$

推定するデータに対する尤度関数  $L$  は MNL モデルなどと同様の形になる。

$$\ln L = \sum_{n=1}^N \sum_{i=1}^J \delta_{in} P_{in} \quad (3.173)$$

ただし、 $\delta_{in}$  は個人  $n$  が選択肢  $i$  を選択したとき 1 をとる指示変数である。

Mixed Logit モデルでは、選択確率の導出式において積分系が残るため、解析的に計算ができない。そのため、乱数計算によるシミュレーションアプローチによる選択確率の計算が必要となる。シミュレーション対数尤度  $\ln SL$  は選択確率を用いて次式で表される。

$$\ln SL = \frac{1}{R} \sum_{r=1}^R \sum_{n=1}^N \sum_{i=1}^J \delta_{in} \ln P_n(i) \quad (3.174)$$

$\delta_{in}$  は、個人  $n$  が選択肢  $i$  を選択したとき 1，そうでないとき 0 をとる指示変数である。上記の尤度関数を最大化するようにパラメータを求め、ランダム係数の期待値と標準偏差を得ることができる。

## 参考文献

- [1] 北村隆一, 森川高行, 佐々木邦明, 藤井聡, 山本俊行: 交通行動の分析とモデリング, 技報堂出版, 2002.
- [2] Chu, C.: A paired combinational logit model for travel demand analysis, *Proceedings of Fifth World Conference on Transportation Research*, Vol. 4, pp. 295-309, 2006.
- [3] Daly, A. and Bierlaire, M.: A general and operational representation of generalised extreme value models, *Transportation Research Part B*, vol. 40, pp. 285-305, 2006.
- [4]
- [5]
- [6] Bhat, C.R.: The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models, *Transportation Research Part B: Methodological*, Vol.45, No.7, pp.923-939, 2011.
- [7] Bhat, C.R., Sidharthan, R.: A simulation evaluation of the maximum approximate composite marginal likelihood (MACML) estimator for mixed multinomial probit models, *Transportation Research Part B: Methodological*, Vol.45, No.7, pp.940-953, 2011.

### 3.2.2 離散・連続選択モデル

離散・連続選択モデルは、何を選択するかについて、それに関連する連続量の選択結果に部分的に依存している場合に、離散選択と連続量の選択を相互に関連づけてモデル化する必要がある、との考えに基づくモデルである。

大きく分けると、Tobin(1958)によるTobitモデルを基にしたモデル群と、効用最大化理論に基づくモデル群がある。福田・力石(2012)は、前者を誘導型、後者を構造型のモデルとして両者の特性を整理し、政策評価のように経済理論に整合的であることが求められる場合は構造型のうちキューンタッカー(Kuhn-Tucker, KT)条件に基づくモデル、現象記述や不完全観測下での行動モデルには誘導型のモデルが望ましいとしている。

離散・連続選択モデルでは、離散選択モデルでは扱うことが難しい以下のような問題に適用できるモデルが提案されている。

- 複数の選択肢のうち1つ以上の選択肢の消費量が0になるような場合の消費者需要
- 複数の選択肢の中からいくつかの選択肢を選ぶ選択問題
- 複数の選択肢の中からいくつかの選択肢を選んでそれぞれにある資源量を配分する問題
- $y_1$ がある閾値  $c$  を超えた場合にのみ  $y_1$  が観測されるという選択メカニズムがある場合に、これを無視して単純な推定を行う場合に生じるバイアス（選択バイアス）の補正

(a) 誘導型モデル

**基本モデル：Tobit モデル**

Tobin(1958)によるTobitモデルは、選択される場合とされない場合の説明変数と目的変数の関係を表すプロビットモデルと、選択される場合の説明変数と目的変数の量的な関係を表す回帰モデルを融合するモデルとして提案された。モデルは次式により、 $y_n > 0$  のときに  $y_n$  が観測され、 $y_n \leq 0$  のときに観測されない状況を表し、打ち切り回帰モデルとも呼ばれる。

$$y_n^* = \beta x_n + \epsilon_n$$

$$y_n = \begin{cases} y_n^* & \text{if } y_n^* > 0 \\ 0 & \text{if } y_n^* \leq 0 \end{cases} \quad (3.175)$$

ここで、 $y_n^*$  は個人  $n$  における  $y$  の値、 $y_n$  は個人  $n$  における  $y$  の観測結果、 $x_n$  は個人  $n$  の  $y$  に対する説明変数ベクトル、 $\beta$  は未知パラメータを示す。Tobitモデルは1変量のモデルだが、Amemiya(1974)は、これを多変量に拡張したモデルを提案した。

Tobit(1958)のモデルの発展形として、Heckman(1974,1979)は選択されるかどうかを決定する離散問題 ( $y_{n1}$  とする) と選択される場合の観測量を表す連続問題 ( $y_{n2}$  とする) が異なる関数で表される場合のモデルを提案した。モデルは次式で表され、 $y_{n1} > 0$  のとき  $y_{i2}$  が観測される。

$$\begin{aligned}
 y_{n1}^* &= \beta_1 x_{n1} + \epsilon_{n1} \\
 y_{n2}^* &= \beta_2 x_{n2} + \epsilon_{n2} \\
 y_{i2} &= \begin{cases} y_{n2}^* & \text{if } y_{n1}^* > 0 \\ 0 & \text{if } y_{n1}^* \leq 0 \end{cases}
 \end{aligned} \tag{3.176}$$

Amemiya(1985)はTobitモデルとそのバリエーションをTypeIからTypeVに整理した。Heckman(1979)のモデルはTypeIIに該当する。TypeIIのモデルは離散問題と連続問題が異なる評価関数で表されることが特徴であり、以降のモデルのベースとなっている。離散問題と連続問題は、それぞれの評価関数の誤差項の相関により結びつけられる。2変量正規分布の共分散構造を考慮した条件付き確率を用いて尤度関数を構築することで、相関を反映したパラメータを推定する。

#### 拡張モデル

Tobitモデルの発展の方向性は、連立方程式の数を増やして扱える状況や同時選択の数を増やす試みと、誤差関数の操作により計算性や柔軟性を高める試みに分けられる。前者の試みとしては、TobitモデルのTypeIIIからVへの拡張、Fang(2008)による多肢選択問題への拡張などが挙げられる。

TypeVは、離散問題を表す1つの式と、連続問題を表す2つの式から成り、離散問題の選択結果により連続問題の2つの式のうちいずれを適用するかが決まるモデルである。2つの式の間を行き来することから、内生的スイッチング回帰モデル (Endogenous Switching Regression Model) とも呼ばれる。Maddala(1983)では、以下のような定式化を行っている。

$$\begin{aligned}
 y_{n1}^* &= \beta_1 x_{n1} + \epsilon_{n1} \\
 y_{n2}^* &= \beta_2 x_{n2} + \epsilon_{n2} \\
 y_{n3}^* &= \beta_3 x_{n3} + \epsilon_{n3} \\
 y_{n1} &= \begin{cases} 1 & \text{if } y_{n1}^* > 0 \\ 0 & \text{if } y_{n1}^* \leq 0 \end{cases} \\
 y_{n3} &= \begin{cases} y_{n3}^* & \text{if } y_{n1} = 0 \\ 0 & \text{if } y_{n1} = 1 \end{cases}
 \end{aligned} \tag{3.177}$$

Fang(2008) は、これまで 2 変量に限定されていた Tobit 型のモデルを、Ordered Probit モデルと組み合わせ、ベイズ推定手法を導入することにより多変量に拡張した BMOPT(Bayesian Multivariate Ordered Probit and Tobit) モデルを提案している。定式化は以下のとおりである。

$$\begin{aligned}
 y_{n1}^* &= \beta_1 x_{n1} + \epsilon_{n1} \\
 y_{n2}^* &= \beta_2 x_{n2} + \epsilon_{n2} \\
 y_{n3}^* &= \beta_3 x_{n3} + \epsilon_{n3} \\
 y_{n4}^* &= \beta_4 x_{n4} + \epsilon_{n4} \\
 y_{nj} &= \begin{cases} 0 & \text{if } y_{nj}^* < \alpha_1 \\ 1 & \text{if } \alpha_1 \leq y_{nj}^* < \alpha_2, \text{ for } j = 1, 2 \\ 2 & \text{if } y_{nj}^* > \alpha_2 \end{cases} \quad (3.178) \\
 y_{n3} &= \begin{cases} y_{n3}^* & \text{if } y_{n1} = 1 \text{ or } 2 \\ 0 & \text{if } y_{n1} = 0 \end{cases} \\
 y_{n4} &= \begin{cases} y_{n4}^* & \text{if } y_{n2} = 1 \text{ or } 2 \\ 0 & \text{if } y_{n2} = 0 \end{cases}
 \end{aligned}$$

$\alpha_1$  は Ordered Probit モデルの Cut point を示し、Fang(2008) では  $\alpha_1 = \Phi^{-1}(1/3)$ ,  $\alpha_2 = -\Phi^{-1}(1/3)$  としている。

一方、誤差関数の操作により計算性を高め適用範囲を広げるモデルとして、Lee(1983), Bhat and Eluru(2009) などが挙げられる。Lee(1983) は、多項ロジットモデルを採用し、選択確率に対して正規分布の分布関数の逆関数をとることにより、誤差分布を近似的に扱う方法を提案した。

さらに、Bhat and Eluru(2009) は、コピュラ関数を導入し、誤差関数の分布を線形で対称とする 2 変量正規分布の仮定を緩和し、非対称な分布をもつ誤差関数にも対応できるモデルを提案した。コピュラ関数の特徴は、2 変量の相関関係を保持しつつ、相関の特徴を表す関数を複数の分布形を試した上で設定できることにある。

#### (b) 構造型モデル

次に、効用最大化理論に基づくモデルについて概説する。

個人  $n$  が、 $I$  個の異なる財（離散選択肢）について、式の資源制約の下で効用が最大になるように各財の消費量  $x_i$  を選択すると仮定したとき、個人  $n$  における直接効用関数  $U_n$  は式 3.179 で表される。

$$\begin{aligned} U_n &= f_n(z_{1n}, z_{2n}, \dots, z_{In}) \\ \sum_{i=1}^I p_i z_{in} &= E_n, \quad \forall z_{in} \geq 0 \end{aligned} \quad (3.179)$$

$p_i$  は各財の価格、 $E_n$  は個人  $n$  のもつ資源の総量とする。

各財の需要関数を導く方法として、間接効用関数を定義してロワの恒等式を用いる方法と、KT 条件を用いて最適化問題を解く方法が提案されている。

#### ロワの恒等式を用いる方法

ロワの恒等式を用いて間接効用関数から需要関数を導き出す方法は、効用最大化理論に基づくモデルの初期から主流な手法として Dubin and McFadden(1984) をはじめ多くのモデルに採用されてきた。導出の手順として、まず資源制約下で直接効用関数を最大化した結果、選択肢  $i$  が選ばれたとて得られる間接効用関数を、以下の式で表す。

$$Y_{in} = Y_{in}(p_i, E_n, x_{in}, s_n, \epsilon_{in}) \quad (3.180)$$

$x_{in}$  は選択肢  $i$  の観測される属性、 $s_n$  は個人  $n$  の社会経済属性、 $\epsilon_{in}$  は非観測特性とする。個人  $n$  が間接効用関数が最大となる選択肢  $i$  を選ぶと考えると、選択確率は次式で表される。

$$P_{in} = \Pr[Y_{in}(p_i, E_n, x_{in}, s_n, \epsilon_{in}) > Y_{jn}(p_j, E_n, x_{jn}, s_n, \epsilon_{jn}), j \in I, j \neq i] \quad (3.181)$$

ここで、非観測特性  $\epsilon_{in}$  が加算型として間接効用関数に含まれると仮定すると、間接効用関数は次式のように表される。

$$Y_{in} = Y_{in}(p_i, E_n, x_{in}, s_n) + \epsilon_{in} \quad (3.182)$$

$\epsilon_{in}$  に i.i.d. ガンベル分布を仮定すると、確率  $P_{in}$  は多項ロジットモデルにより与えられる。

財  $z_i$  に対する需要関数は、間接効用関数にロワの恒等式を適用し、次式のように表される。

$$z_{in}^* = -\frac{\partial Y_{in}(p_i, E_n, x_{in}, s_n, \epsilon_{in})/\partial p_i}{\partial Y_{in}(p_i, E_n, x_{in}, s_n, \epsilon_{in})/\partial E_n} \quad (3.183)$$

需要関数を求めるために間接効用関数  $Y_{in}$  を特定する必要があるが、既往研究において様々な提案がなされてきた。Dubin and McFadden(1984) では、間接効用関数を次式のように設定し、

$$Y_{in} = [\alpha E_n + \beta p_i + \gamma x_{in} + \theta s_n] \exp(-\rho p_i) + \epsilon_{in} \quad (3.184)$$

これを用いて需要関数を式により以下のように導出している。

$$z_{in} = -\frac{\beta}{\alpha} + \frac{\rho}{\alpha}(\alpha E_n + \beta p_i + \gamma x_{in} + \theta s_n) \quad (3.185)$$

#### KT 条件を用いる方法

次に、KT 条件を用いる方法を Wales and Woodland(1983) に基づき以下に示す。

ラグランジュ関数は、複数の変数に 1 つ以上の制約が課せられたときに、関数の停留点を求めるために用いられる。式の最大化問題に対して、ラグランジュ関数を次式により定義する。

$$L_{in} = U_{in}(z) - \lambda \left( \sum_{i=1}^I p_i z_{in} - E_n \right) \quad (3.186)$$

$\lambda$  はラグランジュ乗数である。KT 条件 (Kuhn and Tucker, 1951) は、式を解くための必要十分条件を示すものであり、次式で表される。

$$\begin{aligned} U_{in}(z) - \lambda p_i &\leq 0 \leq z_{in}, \quad i = 1, \dots, M \\ p_i^T z_{in} - 1 &\leq 0 \leq \lambda \end{aligned} \quad (3.187)$$

$U_{in}(x)$  は増加関数であるため、消費者は全ての収入を使うと考えられ、この結果、 $\lambda$  は正となり、また少なくとも 1 つの財は必ず消費される。この消費される財を一般性を損なわずに 1 番めの財とすると、式 3.187 から次の式 3.188 のように置くことができる。

$$\lambda = U_{1n}(z)/p_1 \quad (3.188)$$



これらを用いると、効用最大化のための必要十分条件として式 3.189 が導かれる。

$$\begin{aligned} p_1 U_{in}(z) - p_i U_{1n}(z) &\leq 0 \leq z_{in}, \quad i = 2, \dots, M, \\ p_i^T z_{in} &= 1 \end{aligned} \quad (3.189)$$

式 3.189 は次のように解釈できる。まず  $z_{in} > 0$  の場合は  $p_1 U_{in}(x) - p_i U_{1n}(x) = 0$  となり、 $U_{in}(x)/U_{1n}(x) = p_i/p_1$  となるため、財  $i$  と財 1 の間の限界代替比率は無差別関数上にあり、解は価格の比率（資源制約の傾き）と等しくなる。一方、 $i$  が消費されない場合は、限界代替比率は価格の比率より小さくなる。観測できない個人の選好の差を反映するために、個人の選好がランダムに人口上に分布していると仮定し、式 3.190 のように効用関数にランダム項を導入する。

$$U_{in}(z, \epsilon_{in}) = V_{in}(z) + \epsilon_{in}, \quad i = 1, \dots, M, \quad (3.190)$$

これを用いて式 3.189 の  $U_{in}(z)$  を  $V_{in}(z) + \epsilon_{in}$  に置き換えると、次式を得る。

$$\begin{aligned} (p_1 \epsilon_{in} - p_i \epsilon_{1n}) + [p_1 V_{in}(z) - p_i V_{1n}(z)] &\leq 0 \leq z_{in}, \quad i = 2, \dots, M, \\ p_i^T z_{in} &= 1 \end{aligned} \quad (3.191)$$

分析者には  $\epsilon_{in}$  は観測できないため、 $\epsilon_{in}$  の分布を仮定することにより、式 3.191 を用いて需要  $z_{in}$  の分布を計算する。ここでは、 $\epsilon_{in}$  に平均 0、分散  $\Sigma$  の多変量正規分布を仮定する。式 3.191 の左辺は  $\epsilon_{in}$  について線形なので、 $y_{in} = p_1 \epsilon_{in} - p_i \epsilon_{1n}$  を定義し、平均 0 で分散  $\Omega$  の多変量正規分布に従うものとする。資源制約式 3.188 を  $z$  の 1 要素を消去するのに用いることができるため、これを  $z_{1n}$  に適用すると、式 3.191 は次式のように表される。

$$\begin{aligned} y_i - \bar{y}_i(\hat{z}) &\leq 0 \leq z_{in}, \quad i = 2, \dots, M, \\ \hat{z} &= (z_2, \dots, z_M) \end{aligned} \quad (3.192)$$

$M$  個の財すべてが消費されるとすると、密度関数として次式が得られる。

$$\begin{aligned}
 y_i &= \bar{y}_i(\hat{z}), \\
 f(\hat{z}) &= \phi(\hat{y}, \Omega) \text{abs}[J(\hat{z})], \\
 \hat{y} &= (y_2, \dots, y_M)
 \end{aligned} \tag{3.193}$$

$\phi$  は正規分布の密度関数であり、 $J$  は  $y$  から  $z$  への変換に関するヤコブ行列である。最初の 1 財だけが消費される場合は、式 3.192 における  $M - 1$  個すべての財に関する条件は不等号となる。このため、 $z = \mathbf{0}$  となる場合の確率は次式となる。

$$f(\mathbf{0}) = \int_{-\infty}^{\bar{y}_M} \dots \int_{-\infty}^{\bar{y}_2} \phi(\hat{y}, \Omega) dy_2 \dots dy_M \tag{3.194}$$

以上を一般化し、消費される財の個数を  $K$  個とすると、確率密度関数は次式で表すことができる。

$$\begin{aligned}
 f(z_2, \dots, z_K, \mathbf{0}, \dots, \mathbf{0}) &= \\
 \int_{-\infty}^{\bar{y}_M} \dots \int_{-\infty}^{\bar{y}_{K+1}} \phi(y_2, \dots, y_K, y_{K+1}, \dots, y_M, \Omega) \times \text{abs}[J_K(\hat{x})] dy_{K+1} \dots dy_M
 \end{aligned} \tag{3.195}$$

$J_K$  は  $(z_{K+1}, \dots, z_M) = \mathbf{0}$  のとき、 $(y_2, \dots, y_M)$  を  $(z_2, \dots, z_M)$  に変換するためのヤコブ行列である。式 3.196 は  $(y_2, \dots, y_M)$  から  $(z_2, \dots, z_M)$  への変換が 1 対 1 のとき有効となる。変換が 1 対 1 でない場合は、1 対 1 になるように範囲を分割して積分を行う必要がある。

$M!/K!(M-K)!$  個の可能な消費パターンと、これに対応する確率密度関数が存在する。なお、 $z_1 > \mathbf{0}$  の仮定を置いているため、 $z_1 = \mathbf{0}$  の場合は消費量が 0 でない財を 1 番めの財に置き換えることで密度関数を得る。

Wales and Woodland(1983) は以上の提案により、ロワの恒等式を用いた場合には 1 つの財しか選択できなかつたのに対して、2 つ以上の、かつ全てではない選択肢を、同時に選択して資源を配分するというモデルに、離散・連続選択モデルを拡張した。

#### KT 条件を用いる方法：MDCEV モデル

Bhat(2005,2008) は、以上の Wales and Woodland(1983) のモデル、及びこれを発展させた Kim et al. (2002) のモデルを受けて、ガンベル分布を仮定した Multiple Discrete-Continuous Extreme Value (MDCEV) モデル

を提案した. MDCEV モデルはロジットモデルの自然な拡張と見なすことができ, 離散選択モデルにおいてロジットモデルから Nested logit model や mixed logit model 等への拡張を行う場合と同様な拡張を, 離散・連続選択モデルにおいて行うことが可能となった. MDCEV モデルは, ロジットモデルと同様クローズドフォームのモデルであり, 計算性が著しく高いことが特徴である. 同時に, ロジットモデルと同様, 選択肢間の独立を仮定しており, モデルの大きな制約となっている.

MDCEV モデルの導出過程を以下に示す.

MDCEV モデルでは, 効用関数を次式で表す.

$$U(z) = \sum_i \frac{\gamma_i}{\alpha_i} [\exp(\beta x_i + \epsilon_i)] \cdot \left\{ \left( \frac{z_i}{\gamma_i} + 1 \right)^{\alpha_i} - 1 \right\} \quad (3.196)$$

式中の  $\exp(\beta x_i + \epsilon_i)$  は資源の消費が0の点から1ユニット消費する場合の限界効用に相当する.  $x_i$  は選択肢固有属性や個人属性などの説明変数,  $\beta$  はパラメータである.  $\epsilon_i$  は観測されない要因の影響を表す確率項である.  $\alpha_i$ ,  $\gamma_i$  はともに消費量の増加に伴う効用の低減を表すパラメータであり, さらに  $\gamma_i$  は無差別曲線の位置及び端点解の有無を規定する. 分析者から見ると, 個人は式 (3.196) に示すランダム効用関数を, 式 (3.197) の資源制約の下で最大化する.

$$\sum_{i=1}^I p_i z_i = E \quad (3.197)$$

この最適解を, Wales and Woodland(1983) の手法と同様に, ラグランジュ関数を定義し KT 条件を適用することにより求める. ラグランジュ関数は次式で表される.

$$L = \sum_i [\exp(\beta x_i + \epsilon_i)] (z_i + \gamma_i)^{\alpha_i} - \lambda \left[ \sum_{i=1}^I z_i - E \right] \quad (3.198)$$

これに対する KT 条件は次式となる.

$$\begin{aligned} [\exp(\beta x_i + \epsilon_i)] \alpha_i (z_i^* + \gamma_i)^{\alpha_i - 1} - \lambda &= 0 & \text{if } z_i^* > 0, i = 1, \dots, I \\ [\exp(\beta x_i + \epsilon_i)] \alpha_i (z_i^* + \gamma_i)^{\alpha_i - 1} - \lambda &< 0 & \text{if } z_i^* = 0, i = 1, \dots, I \end{aligned} \quad (3.199)$$

式 3.196 の資源制約条件を用いて次数を減らし、必ず消費される財を 1 番めの財とすると、キューンタッカー条件は以下のように表される。

$$\begin{aligned} V_i + \epsilon_i &= V_1 + \epsilon_1, \text{ if } z_i^* > 0, i = 2, \dots, I, \\ V_i + \epsilon_i &< V_1 + \epsilon_1, \text{ if } z_i^* = 0, i = 2, \dots, I, \text{ where} \\ V_i &= \beta x_i + \ln \alpha_i + (\ln \alpha_i - 1) \ln(z_i^* + \gamma_i), i = 1, \dots, I \end{aligned} \quad (3.200)$$

$\epsilon_i$  について、standard extreme value distribution を仮定し、かつ  $z_i$  から独立で、選択肢間での相関のない分布であることを仮定する。このとき、 $I$  個の選択肢の中から  $M$  個を選び、それぞれに  $z_2$  から  $z_M$  の資源を配分する場合の選択確率は、次式で表される。

$$\begin{aligned} &P(z_2^*, z_3^*, \dots, z_M^*, 0, 0, \dots, 0) | \epsilon_1 \\ &= \left\{ \left( \prod_{j=2}^M g(V_1 - V_j + \epsilon_1) \right)^{|J|} \right\} \times \left\{ \prod_{s=M+1}^I G(V_1 - V_s + \epsilon_1) \right\} \end{aligned}$$

$g$  は standard extreme value density function,  $G$  は standard extreme value distribution とし、 $J$  は次式の要素をもつヤコビアンとする。

$$J_{jh} = \frac{\partial[V_1 - V_{j+1} + \epsilon_1]}{\partial z_{h+1}^*}, \quad i, h = 1, 2, \dots, M-1 \quad (3.202)$$

式 3.196 に変形を加え、最終的に次式を得る。

$$\begin{aligned} &P(z_1^*, z_2^*, z_3^*, \dots, z_M^*, 0, 0, \dots, 0) = \\ &\left[ \prod_{j=1}^M \left( \frac{1 - \alpha_j}{z_j^* + \gamma_j} \right) \right] \left[ \sum_{j=1}^M \left( \frac{z_j^* + \gamma_j}{1 - \alpha_j} \right) \right] \left[ \frac{\prod_{j=1}^M e^{V_j}}{(\sum_{k=1}^I e^{V_k})^M} \right] \quad (M \text{ 個の選択肢}) \end{aligned} \quad (3.203)$$

$M = 1$  の場合、式 3.203 は通常の MNL モデルとなる。また、2 つの選択肢がありそのうち 1 つが必ず消費される場合を考えると、式 3.203 は Dubin and McFadden(1984) をはじめとする single discrete-continuous model と同形となる。

先述したように, i.i.d. ガンベル分布の仮定は大きな制約となることから, MNLの拡張と同様の拡張が Bhat らによって提案されている.

まず, 不等分散性や誤差相関への対応を考慮したモデルとして, mixed MDCEV(MMDCEV) model が提案された (Bhat,2005). MMDCEV モデルでは, 誤差項  $\epsilon_i$  が, 次の 3 つの独立項に分割される. 最初の項  $\zeta_i$  は, 選択肢間での i.i.d. ガンベル分布を仮定した分布である. 2 つめの項  $\eta'w_i$  は, 誤差項のスケールパラメータの選択肢間での差異を推定するもので,  $w_i$  は  $I$  次の列ベクトルで各行が選択肢を表す.  $\eta$  は平均 0 の正規分布であり, 各要素が  $\omega_i^2$  の分散をもつ.  $\omega$  は多変量正規分布である  $\eta$  の分散共分散行列を特徴づける真のパラメータのベクトルである. 3 つめの項  $\mu'x_i$  は, 選択肢間の観測されない効用要素に関する相関の生成メカニズムを表す.  $x_i$  は  $H$  次の列ベクトルであり, 各行は, 共通の非観測要素をもつ選択肢のグループ  $h$  を表す.  $i$  が所属するグループに該当する行は 1 をとり, 他の行は 0 をとる. ベクトル  $\mu$  は  $H$  次であり, 独立した正規分布で各要素が  $\sigma_h^2$  の分散をもつ.  $\sigma$  は  $\mu$  の分散共分散行列を特徴づけるパラメータのベクトルである. このとき, 選択確率は次式で表される.

$$P(z_2^*, z_3^*, \dots, z_M^*, 0, 0, \dots, 0) = \int_{\eta} \int_{\mu} \left[ \prod_{j=1}^M \left( \frac{1 - \alpha_j}{z_j^* + \gamma_j} \right) \right] \left[ \sum_{j=1}^M \left( \frac{z_j^* + \gamma_j}{1 - \alpha_j} \right) \right] \left[ \frac{\prod_{j=1}^M e^{V_j + \eta'w_j + \mu'x_j}}{\left( \sum_{k=1}^I e^{V_k + \eta'w_k + \mu'x_k} \right)^M} \right] (M-1)! dF(\mu|\sigma) dF(\eta|\omega) \quad (3.204)$$

$F$  は多変量累積正規分布である.

この他, 代表的な拡張モデルとして, multiple discrete-continuous nested extreme value (MDCNEV) model (Pinjari and Bhat, 2010), 及びこれをさらに一般化した GEV 型 (multiple discrete-continuous generalized extreme value (MDCGEV) model, Pinjari, 2011) などが提案されている.

Wales and Woodland(1983) は, KT 条件を用いたモデルと Tobit 型のモデルを比較し, 確率項の扱いに違いがあることを指摘している. KT 条件を用いたモデルの場合は, 個人ごとの需要は効用関数の制約条件付き最大化により得られ, 1 つ以上の選択肢の消費量が 0 の場合を含んでおり, また不確

実性は効用関数のパラメータが人口に対してランダムに分布していると仮定することにより導出される。一方、Tobit 型のモデルでは、各個人が同じ効用関数をもつと仮定し、個人の観測消費量は、効用を最大化する消費量と打ち切り型の分布をもつランダム項の和で表される。観測消費量が 1 つ以上の選択肢に対するゼロ消費を含むことは、打ち切り型分布によって可能になっている。

(c) 推定法

#### 誘導型モデルの推定法

##### 完全情報最尤推定法 (FIML: full information maximum-likelihood)

完全情報推定法 (FIML 法) は、基本的な原理は通常の最尤推定法と同じであり、欠損のあるデータについても欠損のない部分の情報を利用して、最尤推定値を求める手法である。

Heckman(1974) は、TypeII の Tobit モデルにおいて 2 項正規分布を仮定した FIML 法を提案し、この手法が以降のモデルでの推定法の基本形になっている。多変量正規分布の条件付き分布の性質、つまり、

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right) \quad (3.205)$$

のとき

$$a_1 | a_2 = b \sim N \left( \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (b - \mu_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^t \right) \quad (3.206)$$

を利用して、尤度関数  $L$  は式 のように表される。

$$\begin{aligned}
L &= \prod_{n: y_{n2} \leq 0} Pr(y_{n2} \leq 0) \times \prod_{n: y_{n2} > 0} [Pr(y_{n1} | y_{n2} > 0) Pr(y_{n2} > 0)] \\
&= \prod_{n: y_{n2} \leq 0} Pr(y_{n2} \leq 0) \times \prod_{n: y_{n2} > 0} [Pr(y_{n2} > 0 | y_{n1}) Pr(y_{n1})] \\
&= \prod_{n: y_{n2} \leq 0} [1 - \Phi(\beta_2 x_{n2})] \times \tag{3.207}
\end{aligned}$$

$$\begin{aligned}
&\prod_{n: y_{n2} > 0} \left[ \Phi \left( \frac{1}{\sqrt{1 - \rho^2}} \left\{ \beta_2 x_{n2} + \frac{\rho}{\sigma_1} (y_{n1} - \beta_1 x_{n1}) \right\} \right) \right. \\
&\quad \left. \times \frac{1}{\sigma_1} \phi \left( \frac{y_{n1} - \beta_1 x_{n1}}{\sigma_1} \right) \right] \tag{3.208}
\end{aligned}$$

この尤度関数を最大にするパラメータの値をシミュレーション法などにより推定する。

Lee(1983)では、同じFIML法の枠組みを用いて、正規分布でない分布を正規分布に近似する手法を用いることで、正規分布の仮定を緩和している。

#### 二段階推定法

Heckman(1979)の二段階推定法は、負荷の大きい最尤推定の計算が難しかった時代に提案され、広く用いられてきた推定法である。これは、式3.208において、 $y_2 > 0$  (つまり  $y_1$  が観測されている) という条件下での  $y_1$  の期待値を次式で表現できることを利用する。

$$\begin{aligned}
E(y_{n1} | y_{n2} > 0) &= E(\beta_1 x_{n1} | y_{n2} > 0) + E(\epsilon_{n1} | y_{n2} > 0) \\
&= \beta_1 x_{n1} + E(\epsilon_{n1} | \epsilon_{n2} > -\beta_2 x_{n2}) \\
&= \beta_1 x_{n1} + (\rho \sigma_1) \frac{\phi(\beta_2 x_{n2})}{\Phi(\beta_2 x_{n2})} \tag{3.209}
\end{aligned}$$

具体的な手順を以下に示す。

#### ステップ1

「 $y_{1n}$  が観測されるかされないか」をダミー変数化して、式を用いて  $\beta_2$  の推定値を計算し、この推定値を用いて式により擬似的な説明変数  $\tilde{\lambda}_n$  を計算する。

$$\tilde{\lambda}_n = \frac{\phi(\beta_2 x_{n2})}{\Phi(\beta_2 x_{n2})} \tag{3.210}$$

## ステップ 2

$y_{1n}$  が観測されている対象者について、式 (3.210) を最小化する最小二乗推定を行い、 $\beta_1$  と  $(\rho\sigma_1)$  を推定する。

$$\sum_{n:y_{n2}>0}^N \left( y_{n1} - \beta_1 x_{n1} + (\rho\sigma_1)\bar{\lambda}_n \right)^2 \quad (3.211)$$

## 構造型モデルの推定法

## ロワの恒等式を用いるモデル

## 完全情報最尤推定法 (FIML: full information maximum-likelihood)

選択肢数の少ない場合には、FIML 法による同時推定が適用できる。この方法では、誤差項  $\epsilon_i$  が正規分布に従うことを仮定する。

## 選択性修正法

より幅広い適用範囲をもつ手法として、選択性修正法が用いられる。三輪ら (2008) に基づき選択性修正法の推定方法を以下に示す。選択性修正法では、まず  $\eta = \mathbf{0}$  と仮定して間接効用関数を効用関数としたロジットモデルを推定する。次に、代替案ごとの選択確率  $P_i$  を算出し、代替案  $i$  の需要関数の誤差項のバイアスを式 (3.212) のように修正し、最小二乗法を適用する。

$$\eta = \sum_{j \neq i} \left\{ \frac{\sigma\sqrt{6}}{\pi} r_j \left( \frac{P_j \ln P_j}{1 - P_j} + \ln P_j \right) \right\} + \nu \quad (3.212)$$

$\sigma$  は全母集団における  $\eta$  の標準偏差、 $r_j$  は  $\eta$  と  $\epsilon_j$  の相関係数、 $\nu$  は選択結果とは独立な誤差項である。上式の右辺第 1 項は選択性修正項と呼ばれ、代替案  $i$  が選択されるという条件下での  $\eta$  の期待値を表す。また、 $\sigma\sqrt{6}r_j = \pi$  が需要関数内の選択性修正項に関する未知パラメータとして推定される。

この手法では、はじめに定義した間接効用関数においては同一であったはずの各未知パラメータが、離散選択モデルと需要関数とにおいて異なった値に推定されるといった、理論上矛盾した結果が得られる。この問題に対処するため、三輪ら (2008) は、離散選択モデルと需要関数との未知パラメータを全てシェアし、全てのパラメータを同時推定する方法を提案している。選択肢別標本抽出データを用いることへの対応として、WESML (weighted exogenous sample maximum likelihood) 法により未知パラメータおよび未



知パラメータの分散共分散行列を推定する。このとき、対数尤度関数は以下の式で表される。

$$\begin{aligned} \ln L(\theta) \\ = \sum_{n=1}^N w(J_n) \{ \ln P(J_n | x_n, \theta) + \ln f(x_{J_n} | x_n, \theta) \} w(J_n) = Q(J_n) / H(J_n) \end{aligned}$$

$N$  はサンプル数、 $J_n$  は個人  $n$  により選ばれた選択肢、 $w(J_n)$  はその母集団シェアとサンプル内シェアの比で表される重み、 $Q(J_n)$  はその母集団シェア、 $H(J_n)$  はそのサンプル内シェア、 $x_n$  は説明変数ベクトル、 $\theta$  はパラメータベクトル、 $x_{J_n}$  は需要量、 $P(\cdot)$  は選択確率を表す関数、 $f(\cdot)$  は正規確率密度関数である。また、WESML 推定量  $\theta^*$  の分散共分散行列  $\Sigma$  は、重み付けを行って得られた推定量の分散共分散行列  $\Lambda$  と、得られた推定量  $\theta^*$  を用いて重み付けを行わずに算出される分散共分散行列  $\Omega$  とを用いて、以下のように計算される。

$$\Sigma = \frac{1}{N} \Omega^{-1} \Lambda \Omega^{-1} \quad (3.214)$$

#### KT 条件を用いたモデル

KT 条件を用いたモデルでは、最尤推定法によりパラメータを推定する。

(1) で示した各財の選択確率を用いて尤度関数を定義する。

$z$  に関する  $N$  個のランダムな観測について、各観測は  $\sum_{K=1}^M M! / K! (M - K)!$  個のパターンのうちただ 1 つの密度関数に対応し、尤度関数は次式で表される。

$$L(\hat{z}_1, \dots, \hat{z}_N) = \prod_{n=1}^N f(\hat{z}_n) \quad (3.215)$$

$z_n$  は  $z$  の  $n$  番目の観測値を示す。効用関数  $U(z, \epsilon)$  のパラメータと共分散行列  $\Sigma$  は最尤推定法により推定できる。ただし多重積分を含むオープンフォームの式であるため計算負荷は大きく、シミュレーション法などの適用が必要となる。

MDCEV モデルは対数尤度関数がクローズドフォームで定義され、推定計算が容易であることが特徴である。また、MDCEV モデルを mixed 型に拡

張した MMDCEV モデルの推定は、対数尤度関数を次式のように設定し、疑似モンテカルロシミュレーションを適用して最尤推定法により行われる。

$$L(\beta, \theta, \gamma, \sigma, \omega) = \sum_{q=1}^Q \left[ \int_{\eta} \int_{\mu} \left[ \prod_{j=1}^M \left( \frac{1 - \alpha_j}{z_j^* + \gamma_j} \right) \right] \left[ \sum_{j=1}^M \left( \frac{z_j^* + \gamma_j}{1 - \alpha_j} \right) \right] \right. \\ \left. \left[ \frac{\prod_{j=1}^M e^{V_j + \eta' w_j + \mu' x_j}}{\left( \sum_{k=1}^I e^{V_k + \eta' w_k + \mu' x_k} \right)^M} (M - 1)! dF(\mu | \sigma) dF(\eta | \omega) \right] \right] \quad (3.216)$$

### 参考文献

- [1] Amemiya, T.: Advanced Econometrics, Harvard University Press, 1985.
- [2] Bhat, C.R.: A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions, Transportation Research Part B: Methodological 39, pp. 679-707, 2005.
- [3] Bhat, C.R.: The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions, Transportation Research Part B: Methodological 42, pp. 274-303, 2008.
- [4] Bhat, C.R., Eluru, N.: A copula-based approach to accommodate residential self-selection effects in travel behavior modeling, Transportation Research Part B: Methodological 43, pp. 749-765, 2009.
- [5] Dubin, J.A., McFadden, D.L.: An econometric analysis of residential electric appliance holdings and consumption. Econometrica 52 (2), pp. 345-362, 1984.
- [6] Fang, H.A.: A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density, Transportation Research Part B: Methodological 42 (1), pp. 736-758, 2008.
- [7] Heckman, J.: Sample selection bias as a specification error, Econometrica 47, pp. 153-161, 1979.
- [8] Lee, L.-F.: Generalized econometric models with selectivity, Econometrica 51, pp. 507-512, 1983.
- [9] Maddala, G. S.: Limited-Dependent and Qualitative Variables in Econometrics, Cambridge University Press, 1983.
- [10] Pinjari, A.R., Bhat, C.: A multiple discrete-continuous nested extreme value (MDCNEV) model: Formulation and application to non-worker activity time-use and timing behavior on weekdays, Transportation Research Part B: Methodological 44, pp. 562-583, 2010.

- [11] Tobin, J.: Estimation of relationships for limited dependent variables, *Econometrica* 26, pp. 24-36, 1958.
- [12] Wales, T.J., Woodland, A.D.: Estimation of consumer demand systems with binding non-negativity constraints, *Journal of Econometrics* 21, pp. 263-285, 1983.
- [13] 星野崇宏, 調査観察データの統計科学, 因果推論・選択バイアス・データ融合, 岩波書店, 2009.
- [14] 福田大輔, 力石真, 離散?連続モデルの研究動向に関するレビュー, 土木計画学研究・講演集 46(CD-ROM), 2012.
- [15] 三輪富生, 山本俊行, 森川高行, 駐車場-駐車時間選択行動への離散-連続モデルの適用と駐車料金施策分析, 都市計画論文集 Vol.43, pp. 34-41, 2008.

### 3.2.3 推定法

モデルのパラメータの推定とは、ある仮定したモデルについて、未知の分布を持つ母集団から抽出された標本データより、そのデータの属する母集団分布の性質を推定することである。具体的には、仮定した理論モデルのパラメータを、観測した標本データに最も適合するように同定することである。ここでは、パラメータ推定の各手法を取りあげる。

#### (a) 的中率による評価

モデルの適合度の評価指標の1つとして、的中率がある。的中率は、個人  $n$  の実際の選択結果が  $i_n$ 、個人  $n$  に関して効用関数の確定項部分が最大となる選択肢として推定される選択結果が  $y_n$  とすると、以下のように表される。

$$\frac{d(y_n = i_n)}{N} \quad (3.217)$$

$d(y_n = i_n)$  は、 $y_n = i_n$  のとき 1 をとる指示変数である。的中率は全サンプルの中で、モデルにより選択結果と選択予測が何パーセント当たったかを示すものである。

的中率は直感的でわかりやすい指標であるが、予測された選択結果が何パーセントの確率で選択されると判断されたかという情報が抜け落ちてしまうため、モデルの評価指標として十分であるとはいえない。そこで、通常は後述する尤度比を用いて、モデルで予測される選択確率を考慮してモデルの評価を行う。

## (b) 尤度関数の設定

離散選択モデルの推定法の中で最も多く用いられる方法が最尤法である。最尤法は得られた標本データについて、最も尤もらしい母集団を母集団の推定値とするという考え方である。最尤法は観測データに対して尤度関数を設定し、尤度関数を最適化することによって未知パラメータを決定する。

ここでは観測データは無作為に抽出されていると仮定する。モデルの推定では、理論モデルが正しいという仮定の下で、観測データに対して最も適切なパラメータ  $\beta$  の探索を行う。

離散選択モデルにおいては、サンプル全体の尤度  $L$  は個々のサンプルのモデルによる理論的選択確率の積で表される。(McFadden, 1974[?])

$$L(\beta) = \prod_{n=1}^N \prod_{i=1}^J P_n(i|\beta)^{d_{in}} \quad (3.218)$$

ただし、 $d_{in}$  は、個人  $n$  が選択肢  $i$  を選択したかどうかのダミー変数である。

実際の計算においては確率の積は非常に小さな値をとるため、計算の便宜上、対数尤度を用いることが多い。

$$\ln L(\beta) = \sum_{n=1}^N \sum_{i=1}^J d_{in} \ln P_n(i|\beta) \quad (3.219)$$

尤度関数が大きいほど、モデルが推定で用いるデータについてあてはまりが良く選択結果をより高い確率で予測できていることを表す。モデルのパラメータ推定では尤度関数が最大となるパラメータ  $\beta$  を求める。

## (c) 尤度比

モデルの適合度を表す指標として、最大化した尤度による尤度比が用いられる。尤度比のベースとして、各選択肢の選択確率が等確率となる「無情報モデル」を考える。効用関数が線形和で表される離散選択モデルの場合、全てのパラメータが 0 であるとき無情報モデルとなる。無情報モデルの尤度関数の値を  $L(\mathbf{0})$  と表し、初期尤度と呼ぶ。無情報モデルは、各選択肢をランダムに選ぶ「当てずっぽう」のモデルと考えることができ、無情報モデルに

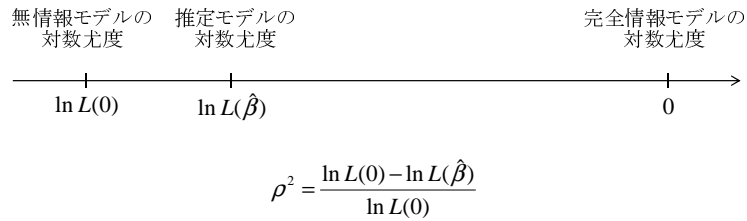


図 3.57 推定モデルの尤度比

比べて推定したモデルがどれだけ説明力が優れているかを表す指標が尤度比である。

無情報モデルの一方で、全てのサンプルの実現選択肢を選択確率 1 として完全に当てるモデルを「完全情報モデル」と呼ぶ。完全情報モデルでは、対数尤度  $\ln L(\beta)$  は 0 となる。尤度関数を最大化するパラメータ  $\beta$  を  $\hat{\beta}$  と表すと、推定したモデルの最大尤度  $L(\hat{\beta})$  は完全情報モデルの対数尤度である 0 から、無情報モデルの対数尤度までの間のいずれかの値をとる。McFadden(1975) は対数尤度の比を用いた適合指標を以下のように提案し、McFadden の決定係数と呼ばれる。

$$\rho^2 = \frac{\ln L(\hat{\beta}) - \ln L(0)}{\ln 1 - \ln L(0)} = \frac{\ln L(0) - \ln L(\hat{\beta})}{\ln L(0)} \quad (3.220)$$

なお、McFadden の決定係数は説明変数を増やすほど増加してしまう欠点がある。McFadden の決定係数をそのまま用いると、パラメータ数を増やすほどモデルの説明力は増加すると解釈され、モデル間の比較などでは都合が良くない。Ben-Akiva and Swait(1986) は、パラメータ数  $K$  について赤池情報量規準 (AIC)  $-\ln 2L(\hat{\beta}) + 2K$  を用いて、自由度調整済み決定係数  $\bar{\rho}^2$  を提案している。

$$\bar{\rho}^2 = \frac{\ln L(0) - (\ln L(\hat{\beta}) - K)}{\ln L(0)} \quad (3.221)$$

## (d) 最適化手法

最尤法では、最適化手法を用いて誘導関数の最大化を行う。

尤度関数  $L$  はパラメータ  $\beta$  に関して大域的に凸であり、 $L$  の導関数が 0 になる解が、尤度が最大となるパラメータである。 $n$  次元のパラメータで表されるモデルの尤度関数の最大化を一般化する。

$$\max_{\beta} L(\beta) \quad (3.222)$$

$$\beta = (\beta_1, \beta_2, \dots, \beta_n) \quad (3.223)$$

また、尤度関数を最大化するパラメータ  $\beta$  を  $\hat{\beta}$  と表すこととする。

最適化の方法として、Newton-Raphson 法、準ニュートン法 (BFGS 法)、最急降下法、2 次導関数を用いた Berndt-Hall-Hall-Hausman (BHHH) 法 (Berndt et al., 1974[?])、また導関数が求められない場合でも適用可能な Nelder-Mead 法 (Nelder and Mead, 1965[?]) などがある。それぞれの数値解法は、収束の早さやヘッセ行列計算の計算量などで一長一短がある。

関数の最適化アルゴリズムとして、以下の方法について紹介する。

- 最急降下法
- Newton-Raphson 法
- BFGS 法 (準ニュートン法)
- BHHH 法
- Nelder-Mead 法

**最急降下法**

最急降下法は、目的関数の勾配ベクトルを用いて解を探索するアルゴリズムである。勾配ベクトルの導出と、更新ステップ幅の決定を交互に繰り返すことにより、解へ近づけていく。

**ステップ 1**

初期解  $\beta_1$  を決定する。

## ステップ 2

$k$  回目ステップにおける探索方向  $\mathbf{s}_k$  を決定する。

$$\mathbf{s}_k = -\nabla L(\beta_k) \quad (3.224)$$

## ステップ 3

ステップ幅  $\alpha_k$  を決定する。

$$\max_{\alpha} L(\beta_k - \alpha_k \mathbf{s}_k) \quad (3.225)$$

## ステップ 4

$\beta_k$  を更新する。

$$\beta_{k+1} = \beta_k - \alpha_k \mathbf{s}_k \quad (3.226)$$

## ステップ 5

$\beta_k$  が収束していなかったら、 $k = k + 1$  として Step2 に戻って計算を繰り返す。

最急降下法はアルゴリズムが単純であるが、解の収束が遅い欠点がある。

## Newton-Raphson 法

Newton-Raphson 法は解を得ようとする定義域において目的関数が微分可能なとき、偏微分値を用いて求解するアルゴリズムである。

## ステップ 1

初期解  $\beta_1$  を決定する。

## ステップ 2

$\beta_k$  に関して尤度関数のヘッセ行列を求める。

$$\nabla^2 L(\beta_k) = \mathbf{A} = \begin{pmatrix} \frac{\partial^2 L}{\partial \beta_1^2} & & \frac{\partial^2 L}{\partial \beta_1 \partial \beta_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 L}{\partial \beta_1 \partial \beta_n} & & \frac{\partial^2 L}{\partial \beta_n^2} \end{pmatrix} \quad (3.227)$$

## ステップ 3

$k$  回目ステップにおける探索方向  $s_k$  を決定する。

$$s_k = -(\nabla^2 L(\beta_k))^{-1} \nabla L(\beta_k) \quad (3.228)$$

## ステップ 4

ステップ幅  $\alpha_k$  を決定する。

$$\max_{\alpha} L(\beta_k - \alpha_k s_k) \quad (3.229)$$

## ステップ 5

$\beta_k$  を更新する。

$$\beta_{k+1} = \beta_k - \alpha_k s_k \quad (3.230)$$

## ステップ 6

$\beta_k$  が収束していなかったら、 $k = k + 1$  として Step2 に戻って計算を繰り返す。

## BFGS 法

BFGS 法は 1970 年に Broyden (1970)[?], Fletcher (1970)[?], Goldfarb (1970)[?], Shanno (1970)[?] によって提案され、4 人の頭文字をとったアルゴリズムである。BFGS 法は Newton-Raphson 法において算出するヘッセ行列を近似行列  $B$  を導入して計算する手法である。ヘッセ行列の計算には多くの計算量が必要であり、ヘッセ行列の計算を回避することにより探索を高速化できる。

## ステップ 1

初期解  $\beta_1$  を決定し、行列  $H_k$  を単位行列としておく。

## ステップ 2

ステップ幅  $\alpha_k$  を決定する。

$$\max_{\alpha} L(\beta_k - \alpha_k H_k \nabla L(\beta_k)) \quad (3.231)$$



## ステップ 3

$\beta_k$  を更新する.

$$\beta_{k+1} = \beta_k - \alpha_k H_k \nabla L(\beta_k) \quad (3.232)$$

## ステップ 4

$s_k = \beta_{k+1} - \beta_k$ ,  $y_k = \nabla L(\beta_{k+1}) - \nabla L(\beta_k)$  として、行列  $H$  を更新する。

$$H_{k+1} = \left[ I - \frac{s_k y_k^T}{s_k^T y_k} \right] H_k \left[ I - \frac{y_k s_k^T}{s_k^T y_k} \right] + \frac{s_k s_k^T}{s_k^T y_k} \quad (3.233)$$

## ステップ 5

$\beta_k$  が収束していなかったら、 $k = k + 1$  として Step2 に戻って計算を繰り返す。

## 最適化手法におけるシミュレーション計算

Mixed Logit モデルやプロビットモデルでは選択確率が積分式で表されるため、シミュレーション法により数値計算を行い、選択確率及び尤度を繰り返し計算により求める最適化手法が必要となる。例としてランダム計数モデルを用いると、パラメータ  $\beta$  が分布  $f(\beta|\theta)$  にしたがう時、個人  $n$  が選択肢  $i$  を選択する確率は一般に以下のようなになる。

$$P_n(i) = \int \frac{\sum_{j=1}^J \exp(V_{jn} + \eta_{jn})}{\exp(V_{in} + \eta_{in})} f(\beta|\theta) d\beta \quad (3.234)$$

上記の積分計算は、一般的に解析的に解くことが困難であり、シミュレーション計算によるアプローチでは、シミュレーションを用いて計算する選択確率を  $SP_n(i)$  とすると、計算の反復回数を  $R$  回とすると、繰り返し計算による試行の平均値をとってシミュレーション尤度とする。(Train, 2003[?])

$$SP_n(i) = \frac{1}{R} \sum_{r=1}^R P_n(i, \beta_r) \quad (3.235)$$

ただし、 $\beta_r$  は確率分布  $f(\beta|\theta)$  から  $r$  回目にサンプリングされたパラメータ  $\beta$  を表す。

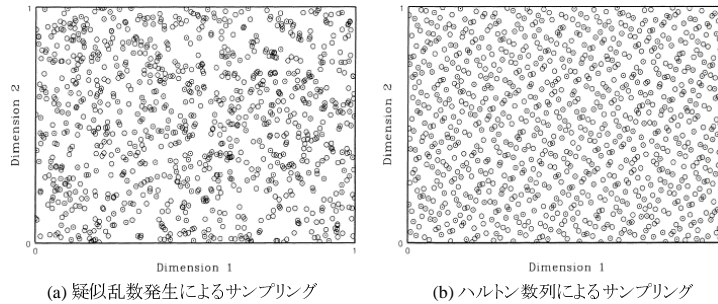


図 3.58 疑似乱数を用いたサンプリングとハルトン数列を用いたサンプリングの比較 (Bhat (2001) より)

シミュレーション計算をより効率的に行う方法として、ハルトン数列による疑似乱数を用いる方法 (Bhat, 2001[?]) がある。ハルトン数列法は通常の乱数の代わりに、分布範囲を均等に網羅しかつ相互に相関などを持たない数列を準乱数として用いる方法である。ハルトン数列法は、ランダムドロー法よりも効率よく分布範囲をカバーできる利点がある。

ハルトン数列は、ベースとなる素数を選び、 $(0,1)$  区間をその数で分割することを繰り返すことにより得る。例えば、素数を 3 とした場合、まず  $(0,1)$  区間を 3 分割する点に対応する数列  $1/3, 2/3$  が得られる。次に、分割した  $(0,1/3), (1/3,2/3), (2/3,1)$  の 3 つの区間について、それぞれ 3 等分する点を順番に加え、 $1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9$  が得られる。シミュレーション計算ではこの数列を  $(0,1)$  区間における準乱数列として用いる。

オープンフォームでない選択確率の計算方法として、EM アルゴリズムを用いた方法 (Train, 2008[?]) やラプラス関数近似 (Harding and Hausman, 2007[?]) などがあり、サンプル数や導出パラメータの安定性などの議論がなされている。(Cherchi, E. and Guevara, 2011[?])

#### (e) ベイズ推定法

ベイズ推定は事前情報とデータからパラメータ値の事後分布の事後分布を更新する方法である。ベイズ推定は、極端にサンプル数が少なく最尤法で計算が安定しない場合や、尤度関数が微分不可能な場合でも推定が可能である

ことが特徴がある。また、長期あるいは複数時点に亘って交通行動を観測する場合、時間の変化に従って行動変化が生じることが考えられ、行動変化の観測や新規データを用いたパラメータの更新にもベイズ推定は適用することができる。

結果の導出は最尤法では推定結果がパラメータの値として求まるが、ベイズ推定ではパラメータ値自体を確率変数と捉え、パラメータ値の分布の形が事後分布として推定される。ベイズ推定では最尤法のような最適化計算を必要としない。尤度関数に積分型の残るオープンフォーム型で表されるプロビットモデルやMixed Logit モデルなどの数値解析的に尤度関数の最大化が難しいモデルでは、計算負荷の大きな関数の最適化計算を回避できる利点は大きい。パラメータの推定値を求める際には、通常は事後分布の母平均を推定結果として用いる。

近年の交通・土木計画学分野のベイズ推定法の適用では、追従挙動モデルパラメータの動的推定（葛西ら, 2012）、day-to-day の交通量推定（Parry and Martin, 2013）、渋滞定着確率モデル（武藤ら, 2010）、OD 推定（山田・羽藤, 2010; Li, 2009; Hazelton, 2001; Maher, 1983）、個人・地域・集団間の異質性を考慮したマルチレベルモデル（古谷・大田, 2010; 力石ら, 2009b）、高速道路交通量の逐次モニタリング（力石, 2009a）、旅行速度推定モデル（Park et al., 2010）、経路選択モデル（河本ら, 2009）、自動車保有-走行距離に関する離散連続モデル（小林ら, 2009）、パーティクルフィルタによる位置推定（薄井ら, 2009; 三谷・羽藤, 2009）、選択肢間の誤差相関を考慮した土地利用モデル（和田ら, 2008）などがみられる。

### 事前分布と事後分布

ベイズ推定は個々のデータに推定において事前分布を仮定して、新たなデータごとに事前分布を更新して事後分布を求める手法である。事前分布とは標本データを得る前のパラメータ分布の情報であり、経験的な数値や過去の分析に基づいて決定される。事後分布とは標本データを得て更新されるパラメータ分布の情報である。ベイズ推定では事後分布の分布形を求めることを目的とする。

行動モデルでは、パラメータを  $\theta$  とすると、尤度関数は各個人  $n$  の選択結

果  $y_n$  に対する選択確率  $P_n(y_n|\theta)$  の積和で表される。

$$L(Y|\theta) = \prod_{n=1}^N P_n(y_n|\theta) \quad (3.236)$$

ただし、 $Y = \{y_1, y_2, \dots, y_n\}$  である。パラメータ  $\theta$  に関する事前分布  $\pi(\theta)$ 、データ  $Y$  が得られて更新されたパラメータ  $\theta$  に関する分布（事後分布）を  $\pi(\theta|x)$  とすると、条件付き確率の乗法の法則より、以下が成り立つ。

$$\pi(\theta|Y)L(Y) = L(Y|\theta)\pi(\theta) \quad (3.237)$$

$L(Y)$  は観測データ  $Y$  そのものについての周辺分布であり、次のように表すことができる。

$$L(Y) = \int L(Y|\theta)\pi(\theta)d\theta \quad (3.238)$$

このとき、式 bayes1 を変形して、ベイズの定理（松原, 2010）が得られる。

$$\pi(\theta|Y) = \frac{L(Y|\theta)\pi(\theta)}{L(Y)} \quad (3.239)$$

ベイズの定理は、事前分布と事後分布の関係を表す式であり、事後分布は事前分布を尤度関数によって重み付けした分布であるといえる。 $L(Y)$  はパラメータ  $\theta$  に依存しない尤度関数  $L(Y|\theta)$  の正規化定数と捉えることができ、 $L(Y)$  が定数であることから、事前分布、事後分布、尤度関数の関係は次のように表すことができる。

$$\pi(\theta|y_t) \propto L(y_t|\theta)\pi(\theta) \quad (3.240)$$

このとき、事後分布  $\pi(\theta|y_t)$  は  $t$  時点までのデータを用いて得られたパラメータ  $\theta$  の確信度であると見なせる。 $t+1$  時点では  $\pi(\theta|y_t)$  は事前分布となり、以下のように逐次パラメータを更新し、パラメータ  $\theta$  を用いたモニタリングを行うことができる。（力石, 2010）

表 3.3 自然な共役事前分布の例

尤度	事前分布	事後分布
二項分布	ベータ分布	ベータ分布
多項分布	ディリクレ分布	ディリクレ分布
正規分布	正規分布 (平均 $\mu$ 、分散は固定)	正規分布
正規分布	正規分布 (平均 $\mu$ 、分散 $\sigma^2$ )	ガンマ分布

$$\pi(\theta|y_t, y_{t+1}) \propto L(y_{t+1}|\theta)\pi(\theta|y_t) \quad (3.241)$$

ベイズの定理を、観測データが入るごとに事後分布を計算するベイズ更新として考えると、 $t$  時点の観測データ  $y_t$  が得られたときの事後分布は

このとき、事後分布を解析的に求めるためには、解析的な積分計算が可能となるように、事前分布と事後分布が同一の分布族となる確率分布を選ぶ必要がある。このような性質を持つ分布を自然共役事前分布と呼ぶ。自然な共役事前分布に関する尤度、事前分布、事後分布の関係を表 3.3 に示す。

尤度分布、事前分布ともに正規分布を仮定するとき、事前分布を平均  $\mu_0$ 、分散共分散行列  $\Sigma_0$  で表される正規分布、尤度関数を平均  $\mu_1$ 、分散共分散行列  $\Sigma_1$  で表される正規分布とすると、ベイズ更新による事後分布のパラメータ  $\hat{\mu}$ 、 $\hat{\Sigma}$  は以下のように表される。

$$\hat{\mu} = \hat{\Sigma}(\Sigma_0^{-1}\mu_0 + \Sigma_1^{-1}\mu_1) \quad (3.242)$$

$$\hat{\Sigma} = (\Sigma_0^{-1} + \Sigma_1^{-1})^{-1} \quad (3.243)$$

自然な共役事前分布を用いたパラメータ推定例として、森地ら (1987) は非集計モデルのパラメータの尤度と事前分布に正規分布を仮定して、集計データと非集計の個人データを統合的に用いてパラメータ推定を行う方法を提案している。さらに屋井ら (1993) は複数の自然な共役事前分布を用いてデータ統合における推定について、比較を行っている。

### ギブスサンプリング

事後分布の計算では積分計算が必要となり計算が解析的には困難であるため、マルコフ連鎖モンテカルロ法 (MCMC 法) によるモンテカルロ計算が多く用いられる。

ギブスサンプリングは、パラメータ間の条件付き確率が既知の場合に適用できるサンプリングアルゴリズムである。ギブスサンプリングは、多次元のパラメータ  $\theta = \{\theta_1, \theta_2, \dots, \theta_k\}$  について、条件付き分布確率にしたがって単一次元ごとにサンプリングを行う方法である。 $i$  番目のサンプルを  $\theta^{(i)} = \{\theta_1^{(i)}, \theta_2^{(i)}, \dots, \theta_k^{(i)}\}$  とするとき、サンプリングは以下のように行う。

#### ステップ 1

初期値  $\theta^{(0)}$  を設定する。 $i = 1$  とする。

#### ステップ 2

$i$  番目のサンプリングにおいて、 $\theta_j^{(i)}$  を条件付き確率分布からサンプリングする。

$$\theta_k^{(i)} \sim f(\theta_k | \theta_1^{(i)}, \dots, \theta_{j-1}^{(i)}, \theta_{j+1}^{(i-1)}, \dots, \theta_k^{(i-1)}) \quad (3.244)$$

#### ステップ 3

$i = i + 1$  として、ステップ 2 から繰り返す。

#### メトロポリス・ヘイスティングス (MH)

MCMC 法を用いた方法の代表例はメトロポリス・ヘイスティングス (Metropolis-Hastings, MH) アルゴリズムである。MH 法では、分布から得られた提案サンプリング値について採択確率を定義し、採択確率にしたがってサンプリング値とするか否かを決定するアルゴリズムである。サンプリングは以下のように行う。

#### ステップ 1

初期値  $\theta^{(0)}$  を設定する。 $i = 1$  とする。

#### ステップ 2

$i$  番目のサンプリングにおいて、提案パラメータ  $\theta^{new(i)}$  を事前分布  $f(\theta^{new(i)} | \theta^{(i)})$  からサンプリングする。

$$\theta^{new(i)} \sim f(\theta^{new(i)} | \theta^{(i)}) \quad (3.245)$$

## ステップ 3

$\theta^{new(i)}$  の採択確率  $\alpha(\theta^{(i-1)}, \theta^{new(i)})$  を求める。

$$\alpha(\theta^{(i-1)}, \theta^{new(i)}) = \min\left(1, \frac{\pi(\theta^{new(i)})f(\theta^{(i-1)}|\theta^{new(i)})}{\pi(\theta^{(i-1)})f(\theta^{new(i)}|\theta^{(i-1)})}\right) \quad (3.246)$$

## ステップ 4

$\theta^{(i)}$  を決定する

$$\theta^{(i)} = \begin{cases} \theta^{new(i)} & (\text{確率 } \alpha \text{ で}) \\ \theta^{(i-1)} & (\text{確率 } 1 - \alpha \text{ で}) \end{cases} \quad (3.247)$$

## 3.2.4 関係性のモデル

## (a) 入れ子の構造が含まれるモデル

本節では、効用間の入れ子の構造が含まれるモデルについて定式化を行い、構造の計算手法を紹介する。

効用間の入れ子の構造が含まれているモデルには、将来の効用を予測して選択を行う動的離散選択モデル、集団との相互作用を表す社会的相互作用モデルが存在する。社会的相互作用モデルについては特定の個人から影響を受けるローカルインタラクションモデル、集団全体から影響を受けるグローバルインタラクションモデルの2つに大別できる。

つまり、動的離散選択モデルは異時点間の関係性を含んだモデル、ローカルインタラクションモデルは個人間の関係性を含んだモデル、グローバルインタラクションモデルは個人と集団の関係性を含んだモデルであると言える。

## 動的離散選択モデル

動的離散選択モデルは、従来の静的な離散選択モデルでは取り扱えない、異時点間の意思決定の変化を記述することができるモデルである。静的離散選択モデルでは、潜在変数を効用関数の差で表すのに対し、動的離散選択モデルでは価値関数の差で表す。ここでの価値関数とは、Bellman 方程式における価値関数を指し、期待将来効用と現在効用の和として個人  $i$  の価値関数は以下のように定義される。

$$V(s_{i,t}) = \max\{U_i(a, s_{i,t}) + \beta \int V(s_{i,t+1})dF(s_{i,t+1} | a, s_{i,t})\} \quad (3.248)$$

この方程式は、将来の効用も含めた全期間を考慮した効用関数である  $\sum_j \beta^j U_i(a, s_{i,t+j})$  を最大化するように選択を行う場合の、dynamic programming(DP) 問題における価値関数を、Bellman の最適性原理 (Bellman,1957) により再帰的に表現した方程式である。この式では図 3.59 のように将来の価値が入れ子の構造となって現在の価値関数に組み込まれているのが特徴である。

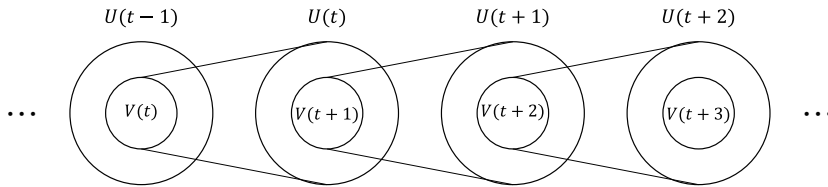


図 3.59 動的離散選択モデルにおける入れ子の構造

ここで、 $U_i(a, s_{i,t})$  は選択枝  $a$  の効用関数、 $s_{i,t}$  は  $t$  期における状態変数、 $dF(s_{i,t+1} | a, s_{i,t})$  は各個人の将来の状態変数に対する信念を表すマルコフ推移確率分布関数、 $\beta$  は価値関数の割引率である。

$s_{i,t}$  のうち観測可能なものを  $x_{i,t}$ 、観測不可能なものを誤差項  $\varepsilon_{i,t}$  とし、誤差項に i.i.d extreme value を仮定すると、選択枝別の価値関数は以下のように表される。

$$v(a, x_{i,t}) = u(a, x_{i,t}) + \varepsilon_{i,t}(a) + \beta \sum_{x_{i,t+1}} \bar{V}(x_{i,t+1}) df(x_{i,t+1} | a, x_{i,t}) \quad (3.249)$$

ここで、 $v(a, x_{i,t})$  は選択枝別価値関数、 $u(a, x_{i,t})$  は選択枝別効用関数、 $\sum \bar{V} df$  は  $t+1$  期に得られる価値関数の期待値である。



この時、各選択肢の選択確率は静的多項ロジットモデルと同様に、以下の式で記述できる。

$$P(a|x_{i,d}, \theta) = \frac{\exp(v(a, x_{i,d}))}{\sum_{j=1}^J \exp(v(j, x_{i,d}))} \quad (3.250)$$

また、誤差項と推移確率の確率密度関数において、条件付き独立性を仮定すると、 $\mathbf{V}$  が一意に定まり、この関数が Bellman 方程式の唯一解となる。

### ローカルインタラクションモデル

社会的相互作用のうち、ローカルインタラクションモデルは特定の個人から受ける影響を考慮していたモデルである。モデルの形は様々であるが、一般的には個人  $i$  が受ける効用の式は以下のような式で表される。

$$U_i = \alpha_i + \beta_i X_i + \sum_{j \neq i} \gamma_j y_j + \varepsilon_i \quad (3.251)$$

ここで、 $\alpha$  は定数項、 $\mathbf{X}_i$  は外生的な説明変数ベクトル、 $\varepsilon_i$  は効用の誤差項である。 $y_j$  は  $i$  が影響を受けている特定の個人  $j$  の影響を表し、選択確率が入る場合や直接  $j$  の効用の確定項  $\mathbf{V}_j$  が入るような場合も存在する。

このモデルは以下の図 3.60 のような、 $i$  が  $j$  に影響を与え、 $j$  が  $i$  に影響を受けているという、効用同士が相互に入れ子になっている構造を持っていることが大きな特徴である。

### グローバルインタラクションモデル

グローバルインタラクションは、個人が集団全体から影響を受けていると仮定したモデルである。ローカルインタラクションとは異なり、集

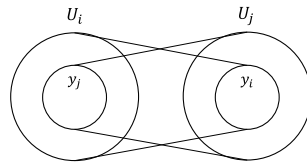


図 3.60 ローカルインタラクションモデルにおける入れ子の構造

団全体の影響の平均を社会的作用として受けている．一般的には以下のような式で表される．

$$U_i = \alpha_i + \beta_i X_i + \gamma_i \frac{\sum_j y_j}{N} + \varepsilon_i \quad (3.252)$$

$N$  は集団の人数を表す．つまり，この式は自分を含む集団全体から受ける効果を平均したものを社会的作用として受けているということになる．このモデルにおける入れ子の構造は図 3.61 のようになる．母集団の数  $N$  が小さいほど個人の行動が集団に与える影響が大きくなる．

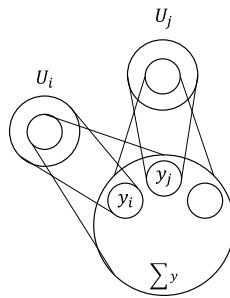


図 3.61 グローバルインタラクションモデルにおける入れ子の構造

#### (b) 計算手法

これまで述べたモデルには効用が入れ子の形で含まれており，通常の最尤推定法では解くことができない．この構造を解く手法として構造推定を用いる．構造推定アプローチは，個人が特定の経済理論モデルの中で最適化行動をとっているという前提の下で，その経済理論を特定化するパラメータを推定

する手法である。代表的な計算手法として、NFXP法、2段階推定法、NPL法の3つを紹介する。

なお、以下では2個人のローカルインタラクションモデルを仮定する。個人  $i \in \{A, B\}$  がサービス  $m$  を利用するか利用しないかの2つの選択肢が与えられている場合、選択する場合の効用  $U_{im}$  を以下のように定式化し、推定を行うものとする。

$$U_{im} = \alpha x_m + \beta_i z_{im} - \lambda y_{-im} + \varepsilon_{im} \quad (3.253)$$

上式において、 $x_m$  は  $m$  の両個人にとって共通の魅力度を表す変数ベクトルで、 $z_{im}$  は個人によって異なる市場の魅力度を表す変数ベクトルである。また、 $y_{-im}$  は相手の選択結果を表すダミー変数である。

### NFXP法

NFXP法 (Nested Fixed Point Method, Rust, 1987) は、繰り返し計算により不動点問題を解くことで入れ子となっている変数を特定し、そのもとでパラメータを更新するという操作をパラメータの値が収束するまで繰り返す手法である。

NFXP法の計算プロセスを図3.62に示す。NFXP法は以下の4つのプロセスからなる。

#### ステップ0:初期化

パラメータの初期値  $\theta^{(0)} = \{\alpha_m^{(0)}, \beta_{im}^{(0)}, \gamma^{(0)}\}$ 、個人  $i \in \{A, B\}$  がサービス  $m$  を利用する確率  $p_{im}$  の初期値  $p_{im}^{(0)}$  を与える。

#### ステップ1:入れ子変数の更新

パラメータ  $\theta^{(n)}$ 、選択確率  $p_{im}^{(k)}$  の下で、選択確率  $p_{im}^{(k+1)}$  を以下のように更新する。

$$p_{im}^{(k+1)} = \Phi(\alpha_m^{(n)} x_m + \beta_{im}^{(n)} z_{im} + \gamma^{(n)} p_{-im}^{(k)}) \quad (3.254)$$

ただし、 $\Phi$  は誤差項の確率分布関数である。

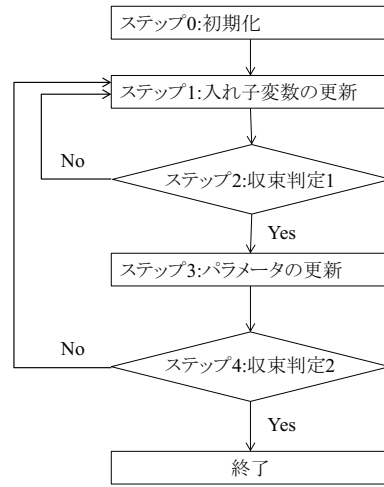


図 3.62 NFXP 法の計算プロセス

**ステップ 2:収束判定 1**

$p_{im}^{(k)}$  と  $p_{im}^{(k+1)}$  を比較し、誤差が十分小さければ次のステップに進む。具体的には閾値を  $\epsilon$  として以下の式を満たせばよい。

$$\left\| p_{im}^{(k+1)} - p_{im}^{(k)} \right\| < \epsilon_p \quad (3.255)$$

上の不等式を満たさない場合は、STEP1 に戻り計算を繰り返す。この繰り返し計算は、以下の不動点問題を解いていることと同義である。

$$p_{im}^* = \Phi(\alpha_m^{(n)} x_m + \beta_{im}^{(n)} z_{im} + \gamma^{(n)} p_{-im}^*) \quad (3.256)$$

**ステップ 3:パラメータの更新**

ステップ 2 までに求めた選択確率の不動点  $p_{im}^*$  および実際の選択結果  $y_{im}$  を用いて尤度関数は以下のように定義できる。

$$\ln L = \sum_m \sum_{i \in \{A, B\}} \{y_{im} \ln(p_{im}^*) + (1 - y_{im}) \ln(1 - p_{im}^*)\} \quad (3.257)$$

この尤度関数を最大化させるようなパラメータ  $\theta^{(n+1)} = \{\alpha_m^{(n+1)}, \beta_{im}^{(n+1)}, \gamma^{(n+1)}\}$  を求める。

#### ステップ 4: 収束判定 2

$\theta^{(n)}$  と  $\theta^{(n+1)}$  を比較し、誤差が十分小さければ計算を終了する。

$$\|\theta^{(n+1)} - \theta^{(n)}\| < \varepsilon_\theta \quad (3.258)$$

### CCP 法

CCP 法 (Conditional Choice Probabilities Method, Hotz, V. J., and R. A. Miller, 1993) は、入れ子の変数を先に決定してしまうことによって不動点問題を解くことを避けた計算手法である。計算プロセスを以下の図 3.63 に示す。

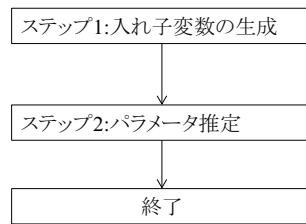


図 3.63 CCP 法の計算プロセス

具体的には以下の 2 つのステップからなる。

#### ステップ 1: 入れ子変数の生成

初めに、入れ子となっている選択確率  $p_{im}^{(1)}$  を観測済み変数  $x_m$ ,  $z_{im}$  により求める。この時、回帰式としてノンパラメトリック回帰を用いることが望ましいが、計算に非常に時間がかかるため、場合によってパラメトリックな回帰や、それら 2 つを組み合わせたセミパラメトリック回帰を用いることもある。

**ステップ 2:パラメータ推定**

ステップ 1 で生成した  $p_{im}^{(1)}$  を元に選択確率を以下のように定義する.

$$p_{im}^{(2)} = \alpha_m x_m + \beta_{im} z_{im} + \gamma_{im} p_{-im}^{(1)} \quad (3.259)$$

この時、対数尤度関数は以下のように定義できる. この対数尤度を最大化させるパラメータを求める.

$$\ln L = \sum_m \sum_{i \in \{A, B\}} \{y_{im} \ln(p_{im}^{(2)}) + (1 - y_{im}) \ln(1 - p_{im}^{(2)})\} \quad (3.260)$$

**NPL 法**

NPL 法 (Nested Pseudo-Likelihood Algorithm, Aguirregabiria and Mira, 2002) は、疑似最尤推定法とも呼ばれ、内生変数を固定した状態における尤度 (疑似尤度) を最大化するようなパラメータを求め、内生変数を更新する操作を収束するまで繰り返す計算手法である. NPL 法の計算プロセスを図 3.64 に示す.

NPL 法は主に以下の 4 つのステップからなる.

**ステップ 0:初期化**

個人  $i \in \{A, B\}$  がサービス  $m$  を利用する確率  $p_{im}$  の初期値  $p_{im}^{(0)}$  を与える. 初期値には一般的に実際の選択結果を与えることが多い.

**ステップ 1:疑似尤度の計算**

選択確率  $p_{im}^{(n)}$  の下での企業  $i$  の選択確率  $\bar{p}_{im}$  は、誤差項の確率分布関数  $\Phi$  を用いて以下のように定義される.

$$\bar{p}_{im} = \Phi(\alpha_m x_m + \beta_{im} z_{im} + \gamma p_{-im}^{(n)}) \quad (3.261)$$

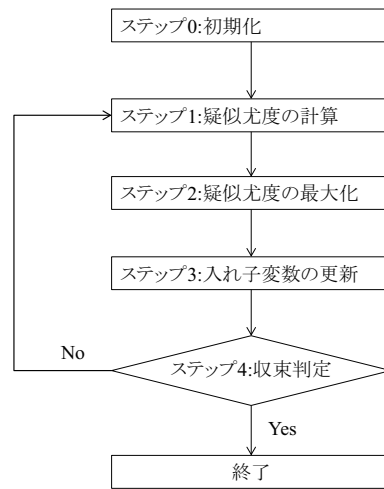


図 3.64 NPL 法の計算プロセス

従って、この時の対数尤度関数 (疑似対数尤度) の値は以下のように定義される。

$$\ln L = \sum_m \sum_{i \in \{A, B\}} \{y_{im} \ln(\bar{p}_{im}) + (1 - y_{im}) \ln(1 - \bar{p}_{im})\} \quad (3.262)$$

#### ステップ 2:疑似尤度の最大化

ステップ 1 で定義した疑似尤度を最大化させるようなパラメータ  $\alpha_m^{(n)}$ ,  $\beta_{im}^{(n)}$ ,  $\gamma^{(n)}$  を最適化計算により推定する。

#### ステップ 3:入れ子変数の更新

ステップ 2 で得られたパラメータ及び  $p_{im}^{(n)}$  から、新しい選択確率  $p_{im}^{(n+1)}$  を以下の式より計算する。

$$p_{im}^{(n+1)} = \Phi(\alpha_m^{(n)} x_m + \beta_{im}^{(n)} z_{im} + \gamma^{(n)} p_{-im}^{(n)}) \quad (3.263)$$

#### ステップ 4:収束判定

$p_{im}^{(n)}$  と新しく出た選択確率  $p_{im}^{(n+1)}$  を比較し、誤差が十分小さければそこで処理を終了する.

$$\|p_{im}^{(n+1)} - p_{im}^{(n)}\| < \varepsilon \quad (3.264)$$

すなわち、閾値を  $\varepsilon$  と置いた時、上式が成り立てば十分収束したとみなし、処理を終了する.

一方誤差が閾値より大きい場合には、 $p_{im}^{(n+1)}$  を新しい選択確率とし、ステップ 1～ステップ 4 を繰り返す.

### 参考文献

- [1] 今村晋, 有村俊秀, 片山東: 労働政策の評価: 「構造推定アプローチ」と「実験的アプローチ」, 日本労働研究雑誌, Vol.43, pp.14-21, 2001.
- [2] Rust, J., Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher, *Econometrica*, Vol.55, pp.999-1033, 1987.
- [3] Hotz, V. J., and R. A. Miller, Conditional Choice Probabilities and the Estimation of Dynamic Models, *Review of Economic Studies*, Vol.60, pp.497-529, 1993
- [4] V, Aguirregabiria., P, Mira., Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models

## 3.3 ネットワークモデル

### 3.3.1 利用者均衡配分

#### (a) 利用者均衡配分の基礎概念

効率的な道路政策を進めるために道路投資の効果をいかに評価するかということが重要になってきている. そのためには道路整備前と後で交通量や、その旅行時間を推計することが必要となってくる. また従前は交通量の再現や予測にその焦点が当てられていたが、近年はマルチモーダル、TDM、ITS 施策といったソフト施策の評価が重要になってきている. そこで利用者均衡配分