交通行動の選択とモデリング
第６章　離散選択モデル

６．１　離散選択モデルの導出

６．２　離散選択モデルの推定

6.1 離散選択モデルの導出

効用　　$U_{in} = V_{in} + \varepsilon_{in}$

個人 n が選択肢 i を選択する確率 Pin は、以下のように表せる。

$$P_{in} = Pr[U_{in} > U_{jn},\ \text{for}\ \forall j,\ i \neq j]$$

$$= Pr[V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn},\ \text{for}\ \forall j,\ i \neq j]$$

$$= Pr[\varepsilon_{jn} < \varepsilon_{in} + V_{in} - V_{jn},\ \text{for}\ \forall j,\ i \neq j]$$

6.1 離散選択モデルの導出

誤差項εを(0,μ)パラメータを持つ、独立で同一なガンベル分布に従うと仮定する

<累積分布関数>　$F(\varepsilon) = e^{-e^{-\mu\varepsilon}}$

<確率密度関数>　$f(\varepsilon) = \mu e^{-\mu\varepsilon} \cdot e^{-e^{-\mu\varepsilon}}$

多項ロジットモデルの導出は、以下の四段階で考える

①あるε_{in}に対して、[$\varepsilon_{jn} < \varepsilon_{in} + V_{in} - V_{jn}$]となる確率

②すべての j について、①が成り立つ確率

③あるε_{in}である確率をかける

④すべてのε_{in}について積分

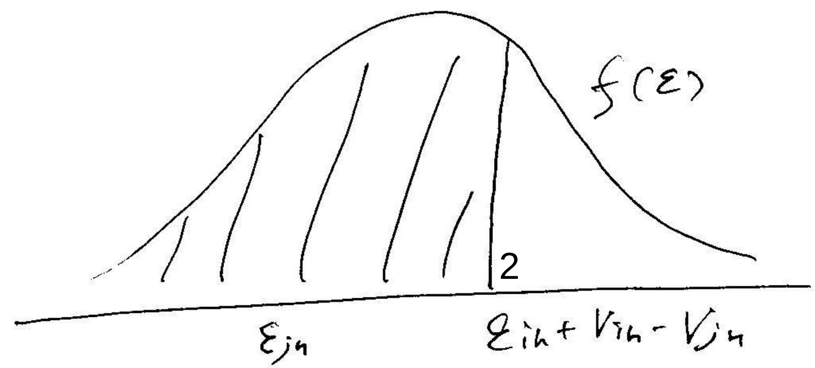
6.1 離散選択モデルの導出

①あるεinに対して、[εjn＜εin＋Vin－Vjn]となる確率

$$\int_{-\infty}^{\varepsilon_{in}+V_{in}-V_{jn}} f(\varepsilon)\,d\varepsilon$$

$$= e^{-e^{-\mu(\varepsilon_{in}+V_{in}-V_{jn})}} - e^{-e^{-(-\infty)}}$$

$$= e^{-e^{-\mu(\varepsilon_{in}+V_{in}-V_{jn})}}$$



例えば、Vin＝Vjn、εin＝2であるとき、

εjn＜2である確率は、図の斜線部の面積

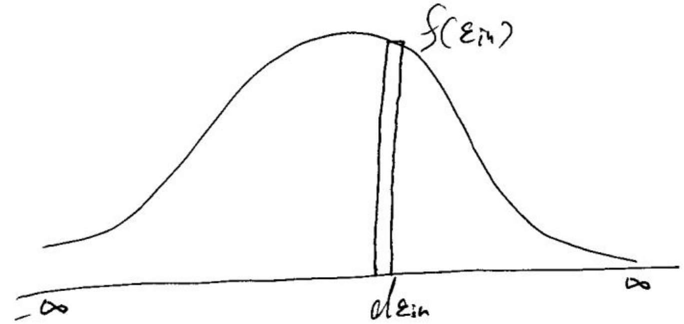
②すべての j について、①が成り立つ確率

$$\prod_{\substack{j\in J_n \\ i\neq j}} \left(e^{-e^{-\mu(\varepsilon_{in}+V_{in}-V_{jn})}} \right)$$

6.1 離散選択モデルの導出

③あるε_{in}である確率をかける

$$f(\varepsilon_{in}) \cdot \prod \left(e^{-e^{-\mu(V_{in}+\varepsilon_{in}-V_{jn})}} \right) = I(\varepsilon_{in})$$



②までは、ε_{in}＝2であるとしたときの確率であるが、ε_{in}＝2 というのも確率的な事象である。

その確率は図の短冊の面積 $f(\varepsilon_{in}) \times d\varepsilon_{in}$ である。

これを積分することにより、すべてのε_{in}について条件を満たす確率が算出される。

④すべてのε_{in}について積分

$$P_{in} = \int_{-\infty}^{\infty} I(\varepsilon_{in})\, d\varepsilon_{in}$$

6.1 離散選択モデルの導出

I (εin) を計算する

$$I(\varepsilon_{in}) = f(\varepsilon_{in}) \cdot \prod_{i \neq j} \left(e^{-e^{-\mu(\varepsilon_{in} + V_{in} - V_{jn})}} \right)$$

$$= \underline{\mu e^{-\mu \varepsilon_{in}} \cdot e^{-e^{-\mu \varepsilon_{in}}}} \cdot \prod_{i \neq j} \left(e^{-e^{-\mu(\varepsilon_{in} + V_{in} - V_{jn})}} \right)$$

$$= \mu e^{-\mu \varepsilon_{in}} \cdot e^{-e^{\underline{-\mu(\varepsilon_{in} + V_{in} - V_{in})}}} \cdot \prod_{i \neq j} \left(e^{-e^{-\mu(\varepsilon_{in} + V_{in} - V_{jn})}} \right)$$

$$= \mu e^{-\mu \varepsilon_{in}} \cdot \underline{\prod} \left(e^{-e^{-\mu(\varepsilon_{in} + V_{in} - V_{jn})}} \right)$$

$$= \mu e^{-\mu \varepsilon_{in}} \cdot e^{\underline{-e^{-\mu \varepsilon_{in}} \cdot \Sigma e^{\mu V_{jn} - \mu V_{in}}}}$$

6.1 離散選択モデルの導出

$I(\varepsilon_{in})$ を積分し、P_{in}を計算 ⇒ 多項ロジットモデルの導出

$$P_{in} = \int_{-\infty}^{\infty} I(\varepsilon_{in})\, d\varepsilon_{in}$$

$$= \int_{-\infty}^{\infty} \mu e^{-\mu\varepsilon_{in}} \cdot e^{-e^{-\mu\varepsilon_{in}} \cdot \Sigma e^{\mu V_{jn} - \mu V_{in}}}\, d\varepsilon_{in}$$

ここで $e^{-\mu\varepsilon_{in}} = t$ とおくと、

$$-\mu e^{-\mu\varepsilon_{in}}\, d\varepsilon_{in} = dt$$

$$d\varepsilon_{in} = \frac{dt}{-\mu t}$$

ε_{in}	$-\infty$	$\cdots$	∞
t	∞	$\cdots$	0

$$= \int_{\infty}^{0} \mu t \cdot e^{-t \cdot \Sigma e^{\mu V_{jn} - \mu V_{in}}}\, \frac{dt}{-\mu t}$$

$$= \int_{0}^{\infty} e^{-t\Sigma e^{\mu V_{jn} - \mu V_{in}}}\, dt$$

$$= \left[\frac{e^{-t\Sigma e^{\mu V_{jn} - \mu V_{in}}}}{-\Sigma e^{\mu V_{jn} - \mu V_{in}}} \right]_0^{\infty}$$

$$= \left(\frac{e^{-\infty\Sigma e^{\mu V_{jn} - \mu V_{in}}}}{-\Sigma e^{\mu V_{jn} - \mu V_{in}}} \right) - \left(\frac{e^{0}}{-\Sigma e^{\mu V_{jn} - \mu V_{in}}} \right)$$

$$= \frac{1}{\Sigma e^{\mu V_{jn} - \mu V_{in}}}$$

$$= \frac{e^{\mu V_{in}}}{\Sigma e^{\mu V_{jn}}}$$

（通常、このμは1として考える）

6.2 離散選択モデルの推定

・モデル分析において、説明変数の関数形を決めたり、誤差項に用いる説明変数を特定の分布系に仮定したりすることを、モデルの特定化と言う。

・離散選択モデルにおける特定化は、それぞれの選択肢の効用関数にどのような説明変数を用いるかを決めること。

特定化の例（自動車：A, 鉄道：R, バス：B）

$$V_{An} = \beta_1 + \beta_3 t_{An} + \beta_4 c_{An} + \beta_5 l_n$$

$$V_{Rn} = \beta_2 + \beta_3 t_{Rn} + \beta_4 c_{Rn}$$

$$V_{Bn} = \beta_3 t_{Bn} + \beta_4 c_{Bn} + \beta_6 f_n$$

（t：所要時間、c：費用、l：免許の有無、f：性別）

・定数項：説明変数では表せなかった効用のうち、全個人で共有する値（選択肢固有定数）
・Level Of Service変数：各選択肢独自のサービスレベルを表す変数
・社会経済変数：意思決定者の属性やトリップの属性を表す

・共通係数 ⇔ 選択肢固有係数（例えば同じ1分が各選択肢の効用に与える影響が異なる場合）

6.2 離散選択モデルの推定

○最尤法によるモデルの推定

・モデルの推定:

理論モデルに含まれる未知パラメータを、実験や観測データを最もよく再現するように同定すること

・最尤推定法(maximum likelihood estimation;MLE)

行動を表す理論モデルが正しいとの仮定の下で、

観測されたデータが得られる尤もらしさが最大になるようにパラメータを定めること

＜尤度関数＞

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

＜対数尤度関数＞

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

6.2 離散選択モデルの推定

〇パラメータ推定値の統計的性質と検定

・推定されたパラメータは、サンプルによって変化する確率変数である

・サンプル数が十分に大きくなれば、統計学的に望ましい性質を持つことがわかっている

最もよく行われる検定は、それぞれの係数が0から有意に離れているか（各説明変数が効用値に影響を与えているか）どうかの検定

⇒　各係数の「ｔ値」の絶対値が1.96以上であれば、十分（有意水準5％で）0から離れているということができる。

6.2 離散選択モデルの推定

〇モデルの適合度

・尤度比(likelihood ratio)を用いる。

無情報モデル尤度L(0)と最大尤度L(β*)の比。（対数尤度の場合は差）

尤度比指標(likelihood ratio index),McFaddenの決定係数

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

この決定係数は説明変数を増やせば必ず指標値が増加するという欠点がある。

これを克服したものが、以下の自由度調整済み決定係数である。

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

0.2程度あれば良い

（0.15でも許せる）

K：モデルに含まれる未知パラメータの数