交通計画　講義ノート

5.　将来の予測　その1

東京工業大学
総合理工学研究科
土木・環境工学科
教授　屋井鉄雄

注意：研究室のHPから資料をダウンロードすること

5．1　ランダム効用理論による交通需要モデル

（1）古典的な効用理論のはなし

○効用の尺度：ユータイル　　図解黒板

限界効用の逓減法則
基数効用案数
序数効用関数

○効用最大化問題：　　図解黒板

max（U）

U＝$U_{映画}$＋$U_{食事}$　：総効用（直接効用関数）

t≧$t_{映画}$＋$t_{食事}$　：時間制約

c≧$c_{映画}$＋$c_{食事}$　：予算制約

→最適化問題（線形関数なら線形計画問題）

直接効用関数を制約条件の下で最大化し，その点の値を効用関数に代入することで，間接効用関数が得られる．

（2）ランダム効用理論，非集計モデルの基礎知識

○ランダム効用理論に基づく離散選択モデル

・ダニエル．マクファーデン（2000年のノーベル経済学賞）が，ミクロ経済学理論に基づく定式化を行った理論体系
・効用関数が確定項とランダム項（確率項）の和で表現できると仮定し，離散選択モデルを活用したパラメータ推定を行う

○合理的経済人の仮定

効用最大化の行動，選択にあたって完全情報を持ち，効用差を識別可能で，効用関数に加法性（要因の同時考慮）を仮定すること（この強い経済的合理性の仮定を弱めたモデルや理論が開発されている：大学院レベル）

○連続と離散

連続量と離散量（連続選択と離散選択）（殺虫剤の効果，労働経済学）
間隔尺度と名義（分類）尺度

→したがって，離散選択(Discrete choice)とは，
複数の選択肢集合から1つを選ぶこと
（東京から大阪への旅行を例にすると，「新幹線と飛行機のいずれか一方を選び，両方を同時には選択できない」）

○非集計と集計

交通量(集計: aggregate)と個人の選択（非集計: disaggregate）
生態学的相関の問題

○選択肢の集合(Choice set)

二肢選択（binary choice) と多肢選択(multiple choice)

ランダム効用理論の関連用語

- 選択確率(Choice Probability)
  - 選択肢(Choice Set)
- 離散選択(Discrete Choice)
- 非集計モデル(Disaggregate Model)
  - 集計モデル(Aggregate Model)
- 効用関数(Utility Function)
  - 確率項，確定項
  - 加法的（線形）効用関数
- ガンベル分布(Gumbel)，正規分布(Normal)
- ロジットモデル(Logit Model),
- プロビットモデル(Probit Model)
- 最尤法(Maximum Likelihood Method)
  - 尤度関数(Likelihood Function)

（3）ランダム効用理論の式展開

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

$U_{nj}, j = 1, \ldots, J$ ：個人nが選択肢jから得る効用（ランダム効用）

V_{nj} ：確定項　例：$U=\beta x+\varepsilon$

ε_{nj} ：誤差項（確率項）　$\varepsilon_n = \langle \varepsilon_{ni}, \ldots, \varepsilon_{nJ} \rangle$

このとき，選択肢 i が選択されるのは，

$$U_{ni} > U_{nj} \ \forall j \neq i$$

なる場合のみである

選択確率　(choice probability : *Pni*)

$$\begin{aligned} P_{ni} &= Prob\left(U_{ni} > U_{nj} \ \ \forall j \neq i\right) \\ &= Prob\left(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \ \ \forall j \neq i\right) \\ &= Prob\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \ \ \forall j \neq i\right) \end{aligned}$$

ここで誤差項の密度分布を $f(\varepsilon_n)$ とおくと，

$$\begin{aligned} P_{ni} &= Prob\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \ \ \forall j \neq i\right) \\ &= \int_{\varepsilon} I\left(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \ \ \forall j \neq i\right) f\left(\varepsilon_n\right) d\varepsilon_n \end{aligned}$$

I(･)は括弧内を満たす場合には1，
そうでなければ0となる判別関数である

上記の積分を行うことで，選択確率を計算することができる

ロジットモデルの理論と展開　(Logit Model)

選択肢ごとの誤差項が各々独立で同一のガンベル分布（第1種極値分布）に従うと考える。密度関数は，

$$f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} \exp(-e^{-\varepsilon_{nj}})$$

また，累積密度関数は，

$$F(\varepsilon_{nj}) = \exp(-e^{-\varepsilon_{nj}})$$

この分布の分散は $\pi^2/6$

独立であることから，すべての選択肢に対する累積分布は，
個々の累積分布の積となる

$$P_{ni} \mid \varepsilon_{nj} = \prod_{j \neq i} \exp(-e^{-\varepsilon_{ni}+V_{ni}-V_{nj}})$$

このとき,ε_{ni} は与えられないので，選択確率*Pni*は密度関数によって重みづけられた，すべてのε_{ni} について$P_{ni} \mid \varepsilon_{ni}$ を積分することで得られる

$$P_{ni} = \int \left(\prod_{j \neq i} \exp(-e^{-\varepsilon_{ni}+V_{ni}-V_{nj}}) \right) - e^{-\varepsilon_{ni}} \exp(-e^{-\varepsilon_{ni}}) d\varepsilon_{ni}$$

この積分を計算すると，以下の簡潔な形が導かれる

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

この式は，ロジットモデル（Logit Model）と呼ばれる。特に3つ以上の選択肢を対象にしていることから，多肢選択ロジット(Multinomial Logit)と呼ばれる

ここで効用の確定項に加法的効用関数を仮定すると，

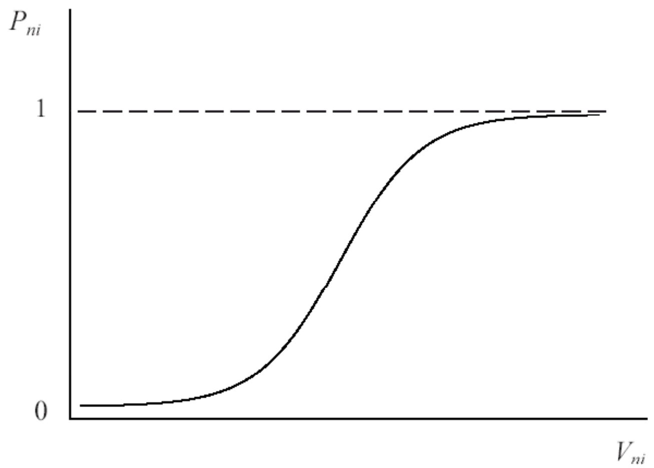
$$V_{nj} = \beta' x_{nj}$$

βはパラメータ，x_{nj} は説明変数である

このとき，以下の式が得られる

$$P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}}$$

選択確率には次の性質がある　$\sum_{i=1}^{J} P_{ni} = \sum_i \exp(V_{ni}) / \sum_j \exp(V_{nj}) = 1$



ロジットモデルのP-Vカーブ

（4）パラメータ推定の理論

○推定に関わる用語

推定(estimation)
推定量(estimator)
推定値(estimate)

○最小2乗推定法：重回帰分析

○最尤推定法：尤度，尤度関数，尤度最大化

福田准教授資料より作成

ロジットモデルのパラメータの推定（1）

最尤法：“未知数を持つ確率”（尤度）の同時生起確率を最大化する未知数を求めるパラメータ推定方法

未知のパラメータ $\boldsymbol{\beta}$, 個人 i の選択結果 y_{ij} (j=1,…,J) と，選択確率 P_{ij} (j=1,…,J) を用いて次のような積をつくる.

$$P_{i1}^{y_{i1}} \cdot P_{i2}^{y_{i2}} \cdot \cdots \cdot P_{iJ}^{y_{iJ}} = \prod_{j=1}^{J} P_{ij}^{y_{ij}}$$

よって，全サンプル n 人の同時生起確率は，

$$L = \prod_{1 \le i \le n} \prod_{1 \le j \le J} P_{ij}^{y_{ij}}$$

によって表される（同時生起確率最大化）

→これが尤度関数

福田准教授資料より作成

ロジットモデルのパラメータの推定（2）

対数尤度関数： $LL = \ln L = \sum_{1 \le i \le n} \sum_{1 \le j \le J} y_{ij} \ln(P_{ij})$

対数尤度最大化：

$$\frac{\partial LL}{\partial \beta_k} = \frac{\partial}{\partial \beta_k}\left[\sum_{1 \le i \le n} \sum_{1 \le j \le J} y_{ij} \ln(P_{ij})\right] = \ldots = \sum_{i=1}^{n} \sum_{j=1}^{J} \left\{\left(y_{ij} - P_{ij}\right) X_{ijk}\right\} = 0 \quad \forall k$$

ベクトル表現：

$$\nabla LL = \begin{pmatrix} \partial LL / \partial \beta_1 \\ \vdots \\ \partial LL / \partial \beta_k \\ \vdots \\ \partial LL / \partial \beta_K \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \sum_{j=1}^{J} \{(y_{ij} - P_{ij}) X_{ij1}\} \\ \vdots \\ \sum_{i=1}^{n} \sum_{j=1}^{J} \{(y_{ij} - P_{ij}) X_{ijk}\} \\ \vdots \\ \sum_{i=1}^{n} \sum_{j=1}^{J} \{(y_{ij} - P_{ij}) X_{ijK}\} \end{pmatrix} = \mathbf{0}$$

すなわち，$\nabla LL(\boldsymbol{\beta}) = \mathbf{0}$ の解 $\hat{\boldsymbol{\beta}}$ が，尤度を最大化する.

福田准教授資料より作成

ロジットモデルのパラメータの推定（3）

対数尤度関数の大域的凹性は保証されている.
すなわち，唯一の最大値が存在する.

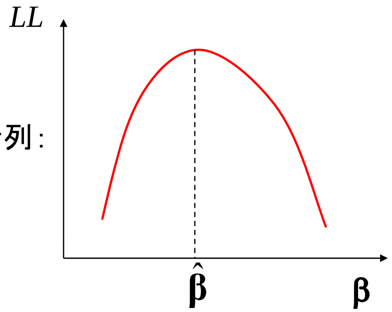
ヘッセ行列のk行l列要素：

$$\frac{\partial^2 LL}{\partial \beta_k \partial \beta_l} = -\sum_{i=1}^{n} \sum_{j=1}^{J} P_{ij} (X_{ijk} - \sum_{j'=1}^{J} P_{ij'} X_{ij'k})(X_{ijl} - \sum_{j'=1}^{J} P_{ij'} X_{ij'l})$$

→ ヘッセ行列：$\nabla^2 LL(\boldsymbol{\beta})$

→パラメータ $\hat{\boldsymbol{\beta}}$ の分散共分散行列：

$$\hat{V} = \left[-\nabla^2 LL(\hat{\boldsymbol{\beta}})\right]^{-1}$$



福田准教授資料より作成

（5）ロジットモデルの有するIIA特性*

＊「無関係な選択肢からの文脈独立特性
（Independence from Irrelevant Alternatives: IIA」

○ ロジットモデルでは，任意の2つの選択肢 j , k に関して，選択確率の比は以下の通り

$$\frac{P_{ij}}{P_{ik}} = \frac{\exp(V_{ij}) \Big/ \sum_s \exp(V_{is})}{\exp(V_{ik}) \Big/ \sum_s \exp(V_{is})} = \exp(V_{ij} - V_{ik})$$

○ 選択確率の相対比は，他の選択肢の影響を受けない

→有名な “青バスー赤バス問題”

福田准教授資料より作成

（6）非集計モデルの適用事例

審議会名	都交審15号（1972年）	運政審7号（1985年）	運政審18号（2000年）
機関分担	集計ロジット	非集計ロジット	非集計ロジット
経路配分	最短経路配分	非集計ロジット	非集計プロビット
経路配分の変数	幹線所要時間	（端末部）所要時間，費用 （幹線部）所要時間，費用，乗換え回数	（端末部）所要時間，費用 （幹線部）所要時間，費用，乗換時間，　混雑費用
トリップ目的		通勤・通学	通勤・通学・私事・業務
ゾーン数		658 （約3.5×3.5km）	1812 （約2×2km）

モデルパラメータの推定結果の例

交通機関選択モデル（多肢選択ロジットモデル）

		通勤の例
時　間	総時間	-0.0263 (-7.9)
費　用	総費用	-0.000584 (-2.0)
乗用車保有率	自動車	0.601 (8.4)
都心ダミー	鉄　道	0.307 (2.6)
定数	自動車	-0.274 (-2.0)
	バ　ス	-1.310 (-7.5)
的　中　率		70.0%
尤　度　比		0.226
サンプル数		2,033
(参考)時間価値		45(円/分)

経路配分モデル（多肢選択プロビットモデル）

		単位	通勤の例
時間	乗車時間	(分)	-0.0943 (-8.09)
	アクセス・イグレス時間	(分)	-0.127 (-11.7)
	アクセス時間	(分)	
	イグレス時間	(分)	
	乗換時間（待ち時間含む）	(分)	-0.112 (-10.7)
費用	総費用	(円)	-0.00200 (-3.98)
混雑指標			-0.00869 (-3.34)
尤　度　比			0.390
サンプル数			1,218

※（　）内は t検定値
通勤・通学・私事・業務目的別に構築

福田准教授資料より作成

（7）非集計ロジットモデルによる集計

非集計モデルで算出される選択確率P_{ij}は，個々人の確率である。これを母集団で集計(aggregation)することでマーケットシェアS_jを求めることができる

①サンプル数え上げ法：　$S_j = \dfrac{\sum_{i=1}^{n} P_{ij}}{n}$

②平均値法：　$S_j = P_j\left(\bar{x}\right)$

③セグメンテーション法：　$S_j = \sum_g \dfrac{n_g}{n} P_j\left(\bar{x}_g\right)$

課題（講義終了時に提出）

（1）講義に登場した推定，推定値，推定量のそれぞれを説明せよ（A41枚で提出）

（2）IIA特性を赤バスー青バス問題によって説明せよ