Discrete choice models with multiplicative error terms

Fosgerau, M., and Bierlaire, M., 2009,
Discrete choice models with multiplicative error terms,
Transportation Research Part.B, Vol.43, Issue.5, pp.494-505.

2009/4/21
論文ゼミ#1
D1　原　祐輔

0. 本日の目次

1. 離散選択モデル(行動モデル)の簡単な復習
   - 例を出しつつ
2. 新しいモデルの定式化の提案と定式化
   - 簡単な例でイメージを
3. 本モデルの持つ特徴
   - 確率分布
   - 弾力性
   - トレードオフ
   - 期待最大化効用
   - 消費者余剰
   - 異質性の取り扱い
4. 推定例
5. まとめ
6. おまけ

1. Introduction

- McFadden and Train (2000)
- 一般化ランダム効用最大化離散選択モデルを導出
- このモデルはconditional indirect utility (CIU)を次式で表す

$$\underline{U^*(z_j, s, \varepsilon_j, v)} \qquad (1)$$

- 観測項 z_j 　非観測項 ε_j
- 観測個人属性 s 　非観測個人属性 v
- データに適用するために，モデルをさらに特定化
- 選択肢 j の確定効用を次式で表す

$$\underline{V_j = V(z_j, s, v)} \qquad (2)$$

- この仮定の下で研究者はパラメータ推定をする

1. Introduction

- 変数z_j ,s ,v はV_j にまとめられるという仮定を置いて，CIUは(3)式で表す

$$U^*(V_j, \varepsilon_j) \qquad (3)$$

- このとき，選択確率は

$$p(i|z,s) = \int P(i|z,s,v)f(v)dv \qquad (4)$$

- f(v) は母集団における v の確率分布

$$p(i|z,s,v) = Pr[U^*(V_i,\varepsilon_i) > U^*(V_j,\varepsilon_j)\forall j] \qquad (5)$$

1. Introduction

- 一般に加法型誤差項で定式化する

$$U^*(V_j, \varepsilon_j) = V_j + \varepsilon_j \qquad (6)$$

  - ここで，ε_jとV_jは独立である
- 個人の非観測異質性v を無視した効用の線形パラメータでの特定化が一般的であり，

$$V_j = V(z_j, s, v) = \beta' x(z_j, s) \qquad (7)$$

- つまり，式(4),(5)は次のように簡単にできる

$$P(i|z,s) = Pr(\beta' x(z_i, s) + \varepsilon_i > \beta' x(z_j, s) + \varepsilon_j \forall j) \quad (8)$$

- 実際にはεの分布において，一定の仮定をおく
- i.i.d. extreme value分布→MNLモデル

ちょっと脱線して復習

- 車の効用,バスの効用が次式で表される

$$U_c = V_c + \varepsilon_c \qquad V_c = \beta_1 T_c + \beta_2 M_c$$
$$U_b = V_b + \varepsilon_b \qquad V_b = \beta_1 T_b + \beta_2 M_b$$

- T: 所要時間，M:金銭的コスト
- たとえばV_c=4, V_b=3のとき，車の方が確定効用が大きい

1. Introduction

- 式(3)から式(6)の流れは自然であるが…

$$U^*(V_j, \varepsilon_j) \qquad (3)$$

$$U^*(V_j, \varepsilon_j) = V_j + \varepsilon_j \qquad (6)$$

- 誤差項が異質的でi.i.d.の仮定を壊していることがありうる.
- これまで多くの研究者が加法型誤差項の分散を観測，非観測，個人属性の関係などを特定化してi.i.d.の仮定を緩和してきた

- この論文では加法型仮定を乗法型に修正

2. Model formation

- 式(9)で与えられるJ個の選択肢の有限集合C上において，一般的な乗法型誤差項を仮定する．
  - $V_j < 0$：効用関数の確定項
  - $\varepsilon_j > 0$：ランダム変数　個人ごとにi.i.d.でV_jとは独立
  - これらの仮定は多くの適用において自然な仮定

- 式(5)の選択確率はこのモデルの下では次式

$$P(i|z,s,v) = Pr(V_i\varepsilon_i \geq V_j\varepsilon_j \forall j) \quad (10)$$

- 乗法型の定式化も，次のような導出を行うことで，加法型の古典的な定式化と関連させることができる

2. Model formation

$$P(i|z,s,v) = Pr(\underline{V_i\varepsilon_i \geq V_j\varepsilon_j})$$
$$= Pr(\underline{-V_i\varepsilon_i \leq -V_j\varepsilon_j})$$
$$= Pr(\underline{\ln(-V_i\varepsilon_i) \leq \ln(-V_j\varepsilon_j)})$$
$$= Pr(\underline{-\ln(-V_i\varepsilon_i) \geq -\ln(-V_j\varepsilon_j)})$$
$$= Pr(\underline{-\ln(-V_i) - \ln(\varepsilon_i) \geq -\ln(-V_j) - \ln(\varepsilon_j)})$$

$$\boxed{-\ln(\varepsilon_j) = \frac{\xi_j}{\lambda} \quad (11)}$$

ξ_j ：ランダム変数
$\lambda > 0$ ：スケールパラメータ

$$P(i|z,s,v) = Pr(\underline{-\ln(-V_i) - \ln(\varepsilon_i) \geq -\ln(-V_j) - \ln(\varepsilon_j)})$$
$$= Pr(-\ln(-V_i) + \underline{\frac{\xi_i}{\lambda}} \geq -\ln(-V_j) + \underline{\frac{\xi_j}{\lambda}})$$
$$= Pr(\underline{-\lambda\ln(-V_i) + \xi_i \geq -\lambda\ln(-V_j) + \xi_j})$$
$$= Pr(\underline{\bar{V}_i} + \xi_i \geq \underline{\bar{V}_j} + \xi_j) \quad (12)$$

$$\boxed{\bar{V}_i = -\lambda\ln(-V_i) \quad (13)}$$

2. Model formation

- 以上より，対数形を用いることで加法型のRUMフレームワークで表現できる

$$P(i|z,s,v) = Pr(\bar{V}_i + \xi_i \geq \bar{V}_j + \xi_j)$$

- 乗法型定式化では正の定数を両辺に掛けることはln (V)に定数を加えることと等価なので，正規化のためにあるパラメータを1か-1にfixするとよい
- 役に立つのは，費用のパラメータを1に正規化すると，他の係数はすべてwilling to pay 指標として解釈することが可能

2. Model formation

- この定式化は非常に一般的であり，これまでの古典的なモデルにおけるMNL, MEV(GEV), mixedなどすべての離散選択モデルで利用できる

- たとえば，MNLの場合

$$P(i|z,s) = \frac{e^{-\lambda \ln(-V_i)}}{\Sigma_{j\in C} e^{-\lambda \ln(-V_j)}} = \frac{(-V_i)^{-\lambda}}{\Sigma_{j\in C}(-V_j)^{-\lambda}} \quad (14)$$

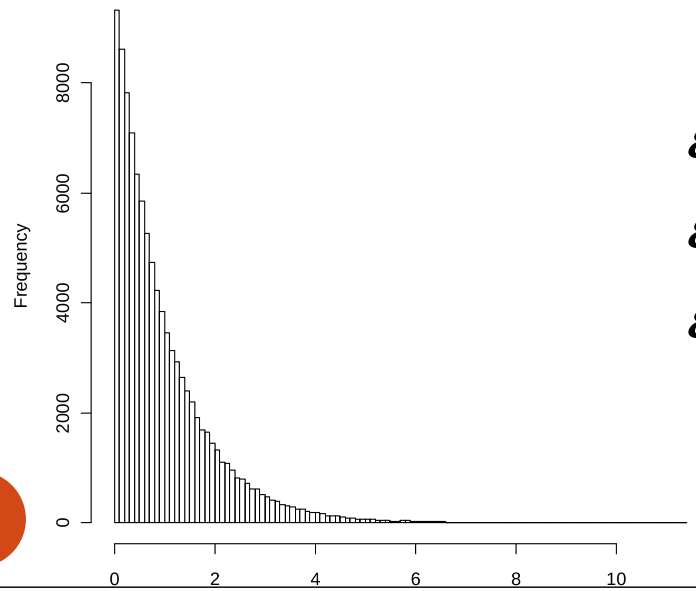
- また，この定式化は効用関数の線形パラメータを置く必要性があまりない（非線形でもOK）
- というか，線形パラメータを置いても，対数を取ったりする時点で非線形になってしまう
- Biogemeなら推定できるよ

ちょっと脱線して数値例でイメージ

	バス	自動車	パラメータ
費用（円）	200	50	−1（fix）
時間（分）	15	10	−12
アクイグ時間（分）	10	2	−15

$V_b =$ $V_c =$

$$\boxed{P(i|z,s,v) = Pr(V_i \varepsilon_i \geq V_j \varepsilon_j \forall j) \quad (10)}$$



たとえば

$\varepsilon_b = 1, \varepsilon_c = 2 \rightarrow$

$\varepsilon_b = 0.5, \varepsilon_c = 1 \rightarrow$

$\varepsilon_b = 0.1, \varepsilon_c = 0.5 \rightarrow$

ちなみにεは指数分布(exponential distribution)に従う（左図参照）理由は後述.

3.1. Distribution

$$-\ln(\varepsilon_j) = \frac{\xi_j}{\lambda} \quad (11)$$

- 式(11)よりε_iのCDFは

$$F_{\varepsilon_i}(x) = 1 - F_{\xi_i}(-\lambda \ln x)$$

- ξ_iがextreme value分布の場合，ξ_iのCDFは

$$F_{\xi_i}(x) = e^{-e^{-x}}$$

- よって，ε_iのCDFは

$$F_{\varepsilon_i}(x) = 1 - e^{-e^{\lambda \ln x}} = 1 - e^{-e^{\ln x^{\lambda}}} = 1 - e^{-x^{\lambda}}$$

- これは指数分布の一般形である

3.2. Elasticities

- 選択肢 i の説明変数xkの直接弾力性は定義より

$$e_{ik} = \frac{\partial P(i)}{\partial x_k}\frac{x_k}{P(i)} = \frac{\partial P(i)}{\partial V_i}\frac{\partial V_i}{\partial x_k}\frac{x_k}{P(i)}$$

- 式(13)を用いて $\bar{V}_i = -\lambda \ln(-V_i)$ (13)

$$e_{ik} = \frac{\partial P(i)}{\partial \bar{V}_i}\frac{\partial \bar{V}_i}{\partial V_i}\frac{\partial V_i}{\partial x_k}\frac{x_k}{P(i)} = -\frac{\lambda}{V_i}\frac{\partial P(i)}{\partial \bar{V}_i}\frac{\partial V_i}{\partial x_k}\frac{x_k}{P(i)}$$

- ここでMNLの場合 $\frac{\partial P(i)}{\partial \bar{V}_i} = P(i)(1 - P(i))$ より

$$e_{ik} = -\frac{\lambda}{V_i}(1 - P(i))\frac{\partial V_i}{\partial x_k}x_k$$

- 同様に交差弾力性は

$$e_{ijk} = -\frac{\lambda}{V_i}(P(j))\frac{\partial V_j}{\partial x_k}x_k$$

3.3. Trade-offs

- トレードオフも加法型と同様に計算できる

$$\frac{\dfrac{\partial U_i}{\partial x_{ik}}}{\dfrac{\partial U_i}{\partial x_{il}}} = \frac{\underline{\dfrac{\partial V_i}{\partial x_{ik}} \cdot \varepsilon_i + V_i \cdot \dfrac{\partial \varepsilon_i}{\partial x_{ik}}}}{\underline{\dfrac{\partial V_i}{\partial x_{il}} \cdot \varepsilon_i + V_i \cdot \dfrac{\partial \varepsilon_i}{\partial x_{il}}}} = \frac{\dfrac{\partial V_i}{\partial x_{ik}}}{\dfrac{\partial V_i}{\partial x_{il}}}$$

$$\frac{\partial \varepsilon_i}{\partial x_{ik}} = \frac{\partial \varepsilon_i}{\partial x_{il}} = 0$$ なので

3.4. Expected maximum utility

- 式(9)における効用の定義下での最大効用は

$$U^* = \max_{i \in C} U_i = \max_{i \in C} V_i \varepsilon_i = \max_{i \in C} V_i e^{-\frac{\xi_i}{\lambda}} \qquad (15)$$

- ξ はGEV分布に従うと仮定すると

$$F(\xi_1, \ldots, \xi_J) = e^{-G(e^{-\xi_1}, \ldots, e^{-\xi_J})} \qquad (16)$$

- G関数とは特徴的な性質を満たした関数

- このとき，期待最大化効用は次式で導出される

$$E[U^*] = -(G^*)^{-\frac{1}{\sigma\lambda}} \Gamma\left(1 + \frac{1}{\sigma\lambda}\right) \qquad (17)$$

3.5. Marshallian consumer surplus

- 消費者余剰は$-V_i$を一般化費用と解釈して導出
- このとき小さな変化dV_iにおける補償変分は選択肢iが選ばれるとき$-dV_i$，それ以外は0
- よって，限界的な変化dV_iにおける補償変分は

$$-P(i)dV_i \qquad (19)$$

- aからbへのV_iの変化の補償変分は

$$-\int_a^b P(i)dV_i \qquad (20)$$

- たとえばP(i)が古典的なMNLで与えられたとき，よく知られたログサム公式になる

$$-\int_a^b P(i)dV_i = -\int_a^b \frac{e^{V_i}}{e^{V_i}+e^{V_j}}dV_i = -\left[\log(e^{V_i}+e^{V_j})\right]_a^b$$

$$= \log(e^a + e^{V_j}) - \log(e^b + e^{V_j})$$

3.6. Heterogeneity of the scale of utility

- 効用を次のように分解することを仮定

$$U^*(V_j,\varepsilon_j)=\tilde{V}(z_j,s)\mu(s,v)\varepsilon_j \quad (21)$$

- 個人の観測異質性s, 非観測異質性は効用のスケールにのみ影響を与える
- 加法型の定式化の下で式(5)(6)を合わせると

$$P(i\mid z,s,v)=\Pr(\tilde{V}(z_i,s)\mu(s,v)+\varepsilon_i>\tilde{V}(z_j,s)\mu(s,v)+\varepsilon_j) \quad (22)$$

- 乗法型の定式化の下で式(5)(9)を合わせると

$$P(i\mid z,s,v)=\Pr(\tilde{V}(z_i,s)\mu(s,v)\varepsilon_i>\tilde{V}(z_j,s)\mu(s,v)\varepsilon_j) \quad (23)$$

$$=\Pr(\tilde{V}(z_i,s)\varepsilon_i>\tilde{V}(z_j,s)\varepsilon_j) \quad (24)$$

- 乗法型の定式化では効用のスケールは確率とは無関係である

4. Empirical applications

1．線形パラメータのロジットモデル

データ：デンマークのSPデータ (車と鉄道)

$$V_i = \lambda(-COST + \beta_1 AE + \beta_2 changes + \beta_3 headway + \beta_4 inVehTime + \beta_5 waiting) \qquad (25)$$

コストのパラメータは-1に固定．λは推定しにいく．
対数に変形すると次式(26)

$$\overline{V_i} = -\lambda \log(COST - \beta_1 AE - \beta_2 changes - \beta_3 headway - \beta_4 inVehTime - \beta_5 waiting) \qquad (26)$$

	加法型誤差項		乗法型誤差項	
説明変数	推定値	t値	推定値	t値
Ae	-2.00	-9.46	-0.672	-11.11
Changes	-36.1	-5.23	-5.22	-3.40
Headway	-0.656	-8.71	-0.224	-10.53
In-veh. time	-1.55	-9.76	-0.782	-11.07
Waiteing time	-1.68	-2.18	-1.06	-5.14
λ	0.0141	9.82	5.37	22.74
observations	3455		3455	
L(0)	-2394.824		-2394.824	
LL	-1970.846		-1799.086	
ρ2	0.177		0.249	
ρ2	0.175		0.246	

5. Concluding remarks

- 今までは加法型の誤差項→乗法型もある！
- 推定もできる！(Biogemeで！)
- アプリオリに効用関数の構造を与えて同じデータで推定したところ，乗法型の方が当てはまりが良い場合が多かった
- しかも，改善度合いは微少ではなく結構大きい！
- 場合によってはmixed logitや潜在クラスにするよりも改善することも
- とにかく推定してみよう

ということで，やってみたが...

- Rだと非線形なので，収束がうまくいかず推定できない・・・