

ロジットモデルの特定と推定

Moshe Ben-Akiva
June 19, 2007
Lecture 2

2010/05/26(水)
M1 柿元淳子

目次

1. 効用関数確定項の特定
2. ロジット推定

選択肢集合

- 選択肢母集合
 - すべての可能な選択肢が母集合に
 - 関連する選択肢に制限される
- 個人の選択肢集合
 - 利用可能性の制約
 - 例) 免許がない場合, 自動車運転の可能性

効用関数のかたち

- $V_{in} = V(Z_{in}, S_n) = V(X_{in})$

Z_{in} —選択肢の属性

S_n —個人属性

- パラメータ線形性の効用関数

$$V_{in} = \beta' X_{in} = \sum_{k=1}^K \beta_k X_{ink}$$

- 選択肢属性と個人属性の相互作用

—たとえば $V_{1n} = \beta_1 + \beta_2 * \underline{Cost_{1n} / Income_n} + \dots$

事例

- 住宅用電話サービスの選択
 - 世帯調査(ペンシルベニア, 1984)
 - 顕示選好(Revealed preference)
 - 434世帯

事例

- 電話サービスと利用可能性

	都心・郊外・一部の 周辺地区	その他の周 辺地区	非都市部
Budget Measured	○	○	○
Standard Measured	○	○	○
Local Flat	○	○	○
Extended Area Flat	×	○	×
Metro Area Flat	×	○	×

5つのサービス

地区によって利用できない選択肢がある。

事例

- 選択肢母集合
 - $C = \{BM, SM, LF, EF, MF\}$
- 特定選択肢集合
 - 都心部・郊外・一部の周辺地区：
 $\{BM, SM, LF, MF\}$
 - 他の周辺地区： C
 - 非都市部： $\{BM, SM, LF\}$

地区によって利用できない選択肢がある。

事例

- 特定表

$$V_{BM} = \beta_1 + \beta_5 \ln(COST(BM))$$

$$V_{SM} = \beta_2 + \beta_5 \ln(COST(SM))$$

$$V_{LF} = \beta_3 + \beta_5 \ln(COST(LF))$$

$$V_{EF} = \beta_4 + \beta_5 \ln(COST(EF))$$

$$V_{MF} = \beta_5 \ln(COST(MF))$$

	β_1	β_2	β_3	β_4	β_5
Budget Measured	1	0	0	0	$\ln(COST(BM))$
Standard Measured	0	1	0	0	$\ln(COST(SM))$
Local Flat	0	0	1	0	$\ln(COST(LF))$
Extended Area Flat	0	0	0	1	$\ln(COST(EF))$
Metro Area Flat	0	0	0	0	$\ln(COST(MF))$

事例

- 特定表

$$V_{BM} = \beta_1 + \beta_5 \ln(COST(BM)) + \beta_6 users$$

$$V_{SM} = \beta_2 + \beta_5 \ln(COST(SM)) + \beta_6 users$$

$$V_{LF} = \beta_3 + \beta_5 \ln(COST(LF)) + \beta_7 MS$$

$$V_{EF} = \beta_4 + \beta_5 \ln(COST(EF))$$

$$V_{MF} = \beta_5 \ln(COST(MF))$$

	β_1	β_2	β_3	β_4	β_5	β_6	β_7
Budget Measured	1	0	0	0	$\ln(COST(BM))$	users	0
Standard Measured	0	1	0	0	$\ln(COST(SM))$	users	0
Local Flat	0	0	1	0	$\ln(COST(LF))$	0	都心か郊外部なら1
Extended Area Flat	0	0	0	1	$\ln(COST(EF))$	0	0
Metro Area Flat	0	0	0	0	$\ln(COST(MF))$	0	0

最大尤度推定①

- 対数尤度関数

$$L(\beta) = \ln L^*(\beta) = \sum_{n=1}^N \ln P(y_n | X_n, \beta) = \sum_{n=1}^N \left(\sum_{j \in C_n} y_{jn} \ln P(j | C_n) \right)$$

nがjを選択すれば, $y_{jn}=1$, そうでなければ $y_{jn}=0$

- ロジットモデル:
$$P(i | C_n) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}}, \mu = 1$$

$$L(\beta) = \sum_{n=1}^N \left(\sum_{j \in C_n} y_{jn} (V_{in} - \ln(\sum_{j \in C_n} e^{V_{jn}})) \right)$$

最大尤度推定②

- 最大尤度推定問題

$$\hat{\beta} = \arg \max_{\beta} L(\beta_1, \beta_2 \dots \beta_K)$$

– ロジットモデルだと

$$L(\beta) = \sum_{n=1}^N \left(\sum_{j \in C_n} y_{jn} (V_{in} - \ln(\sum_{j \in C_n} e^{V_{jn}})) \right)$$

を最大化する β ベクトルを求めればよい

簡単なモデル例 Null model

- $U_i = \varepsilon_i \forall i (\beta = 0)$

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{in}}} = \frac{e^0}{\sum_{j \in C_n} e^0} = \frac{1}{\#C_n}$$

$$L(0) = \sum_n \ln \frac{1}{\#C_n} = -\sum_n \ln(\#C_n)$$

$\#C_n$ は選択肢の数



選択肢の利用可能性に個人差がない。

簡単なモデル例 定数のみ

- $C_n = C$ と仮定

$$U_i = c_i + \varepsilon_i, \forall i$$

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} = \frac{e^{c_i}}{\sum_{j \in C_n} e^{c_j}}$$

$$\ln P(i) = c_i - \ln \sum_j e^{c_j}$$

- サンプル数がNのとき、 N_i (人)が選択肢*i*を選択する。

$$\sum_{n=1}^N y_{in} = N_i$$

- *i*を選択する人すべての対数尤度は

$$L_i = N_i c_i - N_i \ln \sum_j e^{c_j}$$

簡単なモデル例 定数のみ

- 対数尤度の総和

$$L(c) = \sum_i N_i c_i - N \ln \sum_j e^{c_j}$$

- 最大値をとるとき、偏微分の値が0になるので、

$$\frac{\partial L}{\partial c_i} = N_i - N \ln \frac{e^{c_i}}{\sum_j e^{c_j}} = N_i - NP(i) = 0$$

$$P(i) = \frac{N_i}{N}$$

事例の推定結果

Multinomial Logit Model

パラメータ	MNL	
	Value	t値
β BM	-2.46	-7.84
β SM	-1.74	-6.28
β LF	-0.54	-2.57
β EF	-0.74	-1.02
β c	-2.03	-9.47
L0	-560.25	
L	-477.56	