Multiple Choice Models

1) Travel Modes
2) occupational choice
3) Bond ratings (ordered)

UNORDERED models

Random utility, person i and choice j

$$U_{ij} = z_{ij}'\beta + \varepsilon_{ij}$$

if choice j of J possible choices is made then

$$\text{PROB}(U_{ij} > U_{ik}) \text{ for all other } k \neq j.$$

Make model operational by choosing an error dist for e_{ij}.

Probit less popular due to difficulties evaluating multivariate CDF's. logit much more tractable, although there are other disadvantages assoc. with it.

So, $F(\varepsilon_{ij}) = \exp\{-\exp\{-\varepsilon_{ij}\}\}$

which is Gumbel (type 1 extreme value)

Then

$$\text{Prob}(Y_i = j) = \frac{\exp\{\underset{\sim}{z}_{ij}^T \underset{\sim}{\theta}\}}{\sum_{j=1}^{J} \exp\{\underset{\sim}{z}_{ij}^T \underset{\sim}{\theta}\}}$$

which leads to the conditional logit.

Note: $\underset{\sim}{z}_{ij}$ varies by choice $j = 1, \ldots, J$ and individual $i = 1, 2, \ldots, n$

Let $\underset{\sim}{z}_{ij} = [\underset{\sim}{x}_{ij}, \underset{\sim}{w}_i]$ partition the independent variables into 2 sets.

$\underset{\sim}{x}_{ij}$ are <u>attributes</u> that vary ~~across~~ across choices (and possibly individual, though not necessarily)

$\underset{\sim}{w}_i$ are characteristics that vary across individuals but NOT choices.

Partition the parameters.

$$\underset{\sim}{\theta} = (\underset{\sim}{\beta}, \underset{\sim}{\alpha})$$

$$\text{Prob}(Y_i = j) = \frac{\exp(x_{ij}^T \underset{\sim}{\beta} + \underset{\sim}{w}_i \underset{\sim}{\alpha})}{\sum_{j=1}^{J} (\exp(x_{ij}^T \beta) + w_i^T \underset{\sim}{\alpha})}$$

if w_i are same for all choices.

$$= \frac{\exp\{x_{ij}^T \underset{\sim}{\beta}\} \exp\{w_i^T \underset{\sim}{\alpha}\}}{\sum_{j=1}^{J} \exp\{x_{ij}^T \underset{\sim}{\beta}\} \exp\{w_i^T \underset{\sim}{\alpha}\}}$$

$$= \frac{\exp\{x_{ij}^T \underset{\sim}{\beta}\}}{\sum_{j=1}^{J} \exp\{x_{ij}^T \underset{\sim}{\beta}\}}$$

Thus, Terms that do not vary across alternatives fall out of the Model.

Shopping at various Mall's.

Example:

$$U_{i1} = D_{i1}\beta_1 + S_{i1}\beta_2 + \alpha + \gamma I_i + \varepsilon_{i1}$$

$$U_{i2} = D_{i2}\beta_1 + S_{i2}\beta_2 + \alpha + \gamma I_i + \varepsilon_{i2}$$

$$U_{i3} = D_{i3}\beta_1 + S_{i3}\beta_2 + \alpha + \gamma I_i + \varepsilon_{i3}$$

S_{ij} # stores at mall, j

D_{ij} distance to central Bus Dist., j

I_i income

To choose Mall 1

$$U_{i1} - U_{i2} = (D_{i1} - D_{i2})\beta_1 + (S_{i1} - S_{i2})\beta_2 + (\varepsilon_{i1} - \varepsilon_{i2}) > 0$$

$$U_{i1} - U_{i3} = (D_{i1} - D_{i3})\beta_1 + (S_{i1} - S_{i3})\beta_2 + (\varepsilon_{i1} - \varepsilon_{i3}) > 0$$

Random Utility is based on the comparison of alternatives, not the alternatives themselves

If the model is to include individual effects, then it must be modified in some way.

One common way to do this is to create a set of alternative specific indicator vars A_j for each choice and interact these with w_i

This allows the coefficients to differ α_j for each choice.

To avoid perfect collinearity (identify the params) - omit 1 set of interactions

$$
Z_i = \begin{bmatrix} S_{i1} & D_{i1} & 1 & 0 & I_i & 0 \\ S_{i2} & D_{i2} & 0 & 1 & 0 & I_i \\ S_{i3} & D_{i3} & 0 & 0 & 0 & 0 \end{bmatrix}
$$

Then,

$$
\text{Prob}\left[Y_i = j \mid \underset{\sim}{Z_i}\right] = \frac{\exp\left(\begin{array}{c} \text{Stores}_{ij}\beta_1 + \text{Distance}_{ij}\beta_2 \\ + A_1\alpha_1 + A_2\alpha_2 + A_3\alpha_3 \\ + A_1\,\text{Income}_i\gamma_1 + A_2\,\text{Income}_i\gamma_2 \\ + A_3\,\text{Income}_i\gamma_3 \end{array}\right)}{\displaystyle\sum_{j=1}^{3} \exp\left(\begin{array}{c} \text{Stores}_{ij}\beta_1 + \text{Distance}_{ij}\beta_2 \\ A_1\alpha_1 + A_2\alpha_2 + A_3\alpha_3 + \\ A_1\,\text{Inc}_i\gamma_1 + A_2\,\text{Inc}_i\gamma_2 + A_3\,\text{Inc}_i\gamma_3 \end{array}\right)}
$$

with $\alpha_3 = \gamma_3 = 0$

Two special cases

- data consist of characteristics only w_i.

① **Multinomial logit:** Schmidt & Strauss 1975

$n = 1000$

occupational choice

0 = menial, 1 = blue collar

2 = craft, 3 = white collar, 4 = profession

~~BM~~

characteristics:

constant, education, experience, race, gender.

$$\text{Prob}(Y_i = j) = \frac{\exp(w_i^T \alpha_j)}{\sum_{j=0}^{4} \exp(w_i^T \alpha_j)}$$

$j = 0, 1, 2, 3, 4$

MNL

use when data are individual specific

Schmidt & Strauss (1975)

occupation	regressors
0 menial	const
1 ~~2~~ blue collar	ed
2 ~~3~~ craft	exp.
3 ~~4~~ white collar	race
4 ~~5~~ professional	Gender.

$$P(Y_j = j) = \frac{e^{x_i'\beta_j}}{\sum_{k=0}^{4} e^{x_i'\beta_k}} \qquad j = 0, 1, \ldots, 4$$

$J+1 = 5$ choices. ~~since~~ $\beta_j^* = \beta_j + g$ for any g has NO effect on the resulting probabilities. $\therefore$ a normalization is required. The usual 1 is $\beta_0 = 0$

$$P(Y = j) = \frac{e^{x_i'\beta_j}}{1 + \sum_{k=1}^{J} e^{x_i'\beta_k}} \qquad j = 1, 2, 3, 4, \ldots, J$$

$$P(Y = 0) = \frac{1}{1 + \sum_{k=1}^{J} e^{x_i'\beta_k}}$$

Estimation

Newton's method quite effective.

$$\ln(L) = \sum_{i=1}^{T} \ln P(Y_i = 1) + \ln P(Y_i = 2) + \cdots$$

$$+ \ln P(Y_i = 4)$$

or

$$\ln(L) = \sum_{i=1}^{T} \sum_{j=0}^{J} d_{ij} \ln(P(Y_i = j))$$

$$d_{ij} = \begin{cases} 1 & \text{if alternative } j \text{ chosen by } i \\ 0 & \text{otherwise.} \end{cases}$$

✓ $$\frac{\partial \ln L}{\partial \beta_j} = \sum_i (d_{ij} - P_{ij}) x_i \qquad j = 1, \ldots, J$$

$$\frac{\partial^2 l}{\partial \beta_j \partial \beta_l} = -\sum_{i=1}^{T} (P_{ij} [1(j=l) - P_{il}) x_i x_i'$$

$$1(j=l) = \begin{cases} 1 & \text{if } j = l \\ 0 & \text{otherwise.} \end{cases}$$

$$\frac{\partial^2 \ell}{\partial \beta_j \partial \beta_k'} = I_{jk} = \sum_{i=1}^{T} P_{ij} P_{ik} x_i x_i' \qquad j \neq k$$

$$\frac{\partial^2 \ell}{\partial \beta_j \partial \beta_j'} = I_{jj} = -\sum_{i=1}^{T} \left(P_{ij} - P_{ij}^2 \right) x_i x_i'$$

Marginal Effects

$$\delta_j = \frac{\partial P_j}{\partial x_i} = P_j \left[\beta_j - \bar{\beta} \right]$$

May or may not have the same sign as β_j!

Std errors computed using a variant of the δ-Theorem.

Asy Var.

$$= \sum_{l=0}^{J} \sum_{m=0}^{J} \frac{\partial \delta_j}{\partial \beta_l'} \, \text{Asy Cov}(\hat{\beta}_l, \hat{\beta}_m) \frac{\partial \delta_j'}{\partial \beta_m}$$