# 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, if $\varepsilon_{ij} \sim \exp(-e^{-\varepsilon_{ij}})$ Weibull dist

$$\text{Prob}(Y_i = j) = \frac{e^{x_{ij}'\beta_j}}{\sum_{j=1}^{J} e^{x_{ij}'\beta_j}}$$

This leads to conditional logit model.

Note: If $\beta_j^* = \beta_j + q$ then the probabilities are unaffected. So, ~~So~~ we fix this indeterminancy by assuming $\beta_0 = 0$

---

later

$$\text{Prob}(Y_i = j) = \frac{e^{x_{ij}'\beta_j}}{1 + \sum_{k=1}^{J} e^{x_{ij}'\beta_k}}$$

$$\text{Prob}(Y_i = 0) = \frac{1}{1 + \sum_{k=1}^{J} e^{x_i'\beta_k}}$$

There are J log-odds ratios

$$\ln\left[P_{ij}/P_{i0}\right] = x_i'\beta_j$$

normalizing on other probabilities

$$\ln\left[P_{ij}/P_{in}\right] = x_i'(\beta_j - \beta_n)$$

note that the odds ratio DOES NOT depend on any of the other choices. This feature is referred to as the independence of irrelevant alternatives IIA

This is an important restriction to place on human behavior.

P 4:1

Bus

Car F D

Train

4:1

2:1

2:1

Comute.

alone Bus car pool
Choice A, B, C

Normalize on C

$$\frac{\text{Prob}(A)}{\text{Prob}(C)} = e^{x'\beta_A} \qquad \frac{\text{Prob}(B)}{\text{Prob}(C)} = e^{x'\beta_B}$$

$$\text{Prob}(A) = \frac{e^{x'\beta_A}}{1 + e^{x'\beta_A} + e^{x'\beta_B}}.$$

etc.

Now suppose we add another, trivially different alternative

Blue Bus, Red Bus
A, BB, RB, C

Assuming no one cares about Bus color it stands to reason that the underlying Probs should not change. Only Bus ridership is broken in 1/2 between Blue & Red.

This Does NOT happen in Logit.

Adding irrelevant alternatives changes all Probs

$$\cancel{\text{Prob}(A) = \frac{e^{x\beta_A}}{\Sigma}}$$

There is virtually no change in the individual characteristics

$$\therefore \beta_{RB} \simeq \beta_B \simeq \beta_{BB}.$$

The addition of an irrelevant alternative then has the effect

$$\text{Prob}(A) = \frac{e^{x\beta_A}}{1 + e^{x\beta_{RB}} + e^{x\beta_{BB}} + e^{x\beta_A}}$$

$$= \frac{e^{x\beta_A}}{1 + 2e^{x\beta_B} + e^{x\beta_A}}$$

Adding a bus color effectively ~~[illegible]~~ Reducing the Prob of riding alone!

## MNL

use when data are individual specific

Schmidt & Strauss (1975)

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

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

$J+1 = 5$ choices. $\beta_j^* = \beta_j + q$ for any $q$ 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 $$\ln(L) = \sum_{i=1}^{T} \sum_{j=0}^{J} d_{ij} \ln (P(Y_i = j))$$

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

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

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

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

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

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

## Marginal Effects

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

May or may not have the same sign as $\beta_j$!

Std errors computed using variant of the G - Theorem.

Asy Var.

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

# Conditional Logit

Let $F(e_{ij})$ be iid Weibull. and let

$z_{ij} = [\, x_{ij} \;\; w_i \,]$ where $x_{ij}$ varies across individual & choice, and $w_i$ only contains individual specific characteristics.

$x_{ij}$ - attributes of choice

$w_i$ - characteristics

$$P(Y_i = j) = \frac{e^{x_{ij}'\beta + w_i'\alpha}}{\sum_{j=1}^{J} e^{x_{ij}'\beta + w_i'\alpha}}$$

NOTE:

$$\frac{e^{x_{ij}'\beta} \cancel{e^{w_i'\alpha}}}{\sum_{j=1}^{J} e^{x_{ij}'\beta} \cancel{e^{w_i'\alpha}}}$$

To allow individual effects - the model has to be modified. One possibility is to allow $\alpha$ to vary ~~across~~ choices $\alpha_j$.

Choice specific Attributes



## Conditional Logit

$$\text{Prob}(Y_i = j) = \frac{e^{z_{ij}'\beta}}{\sum_{j=1}^{J} e^{z_{ij}'\beta}}$$

same $\beta$ - different $x_{ij}$

In this literature there are J alternatives

Basically same or more except.

Marginal effects

$$\frac{\partial P_j}{\partial x_k} = -P_j P_k \beta \qquad j \neq k$$

$$\frac{\partial P}{\partial x_k} = (P_j - P_j^2)\beta \qquad j = k$$

Every attribute set $x_j$ affects all probabilities

Estimation

$$\ln(L) = \sum_{i=1}^{T} \ln P(Y_1 = 1) + \cdots + \ln P(Y_J = J)$$

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{T} \sum_{j=1}^{J} d_{ij} \left( x_{ij} - \bar{x}_i \right)$$

$$\frac{\partial^2 l}{\partial \beta \partial \beta'} = \sum_{i=1}^{T} \sum_{j=1}^{J} \cancel{P_{ij}}\ P_{ij} \left( x_{ij} - \bar{x}_i \right)\left( x_{ij} - \bar{x}_i \right)'$$

$$\bar{x}_i = \sum_{j=1}^{J} P_{ij} x_{ij}$$