LECTURE / DISCUSSION

Introduction to Logit

Qualitative Choice Situation

Decisionmaker chooses among several alternatives:

- finite number of alternatives
- mutually exclusive alternatives
- exhaustive set of alternatives

General Specification

J_n = choice set for person *n*

z_{in} = attributes of alternative $i \in J_n$ as faced by person *n*

s_n = characteristics of person *n*

β = parameters

Model: $P_{in} = f(z_{in}, z_{jn} \text{ for all } j \text{ in } J_n, s_n; \beta)$

Logit

$$P_{in} = \frac{e^{V_{in}}}{\sum_{j \in J_n} e^{V_{jn}}}$$

where V_{in} is a function of z_{in} , s_n , and β.

Example

Commuter's Choice between Car and Bus for Trip to Work

$$V_c = \alpha C_c + \beta T_c$$

$$V_b = \alpha C_b + \beta T_b$$

where C is cost and T is time. Note $\alpha < 0$ and $\beta < 0$.

$$P_c = \frac{e^{\alpha C_c + \beta T_c}}{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b}}$$

$$P_b = \frac{e^{\alpha C_b + \beta T_b}}{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b}}$$

Properties of Choice Probabilities

1. Probabilities range from zero to one.

If $T_b = \infty$, then

$$\alpha C_b + \beta T_b = -\infty$$

and

$$e^{\alpha C_b + \beta T_b} = 0$$

so

$$P_b = \frac{0}{e^{\alpha C_c + \beta T_c} + 0} = 0$$

If $T_c = \infty$, then

$$\alpha C_c + \beta T_c = -\infty$$

and

$$e^{\alpha C_c + \beta T_c} = 0$$

so

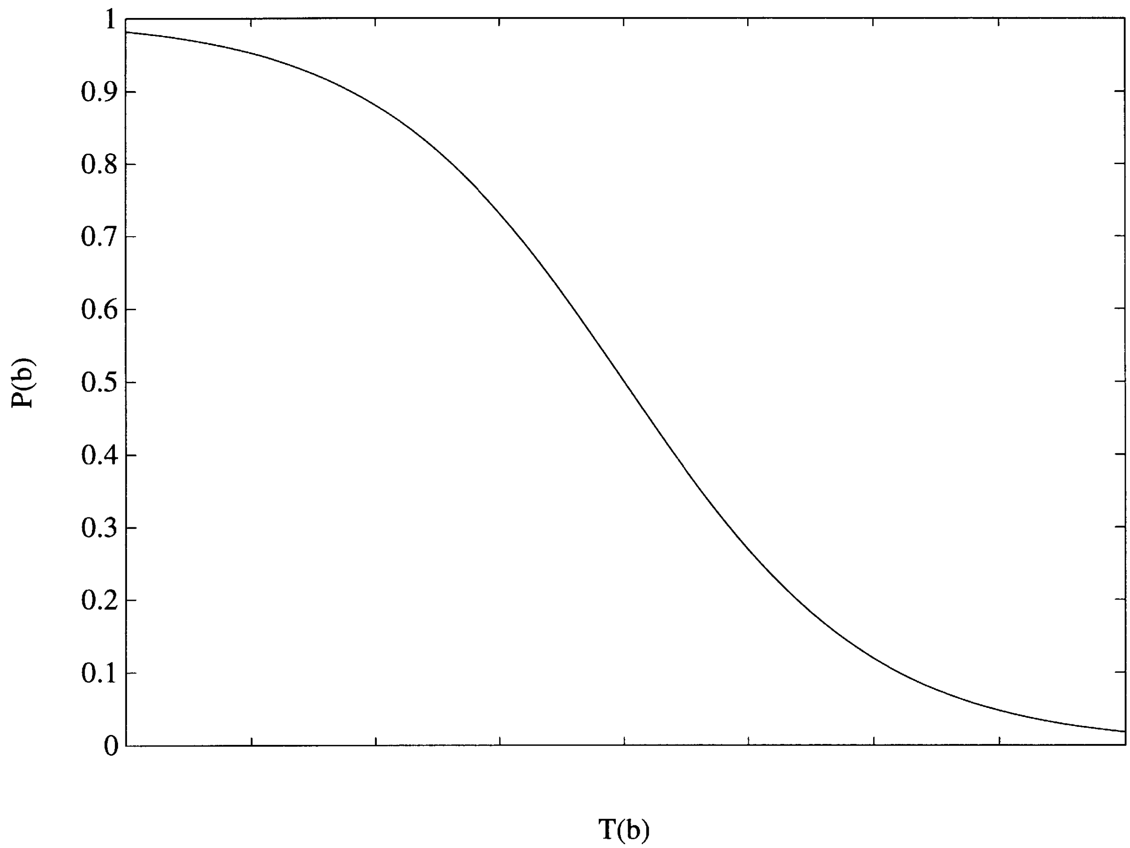
$$P_b = \frac{e^{\alpha C_b + \beta T_b}}{0 + e^{\alpha C_b + \beta T_b}} = 1$$

2. Probabilities sum to one over alternatives.

$$P_c + P_b = \frac{e^{\alpha C_c + \beta T_c}}{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b}} + \frac{e^{\alpha C_b + \beta T_b}}{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b}}$$

$$= \frac{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b}}{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b}} = 1$$

3. Relation between probabilities and explanatory variables is S-shaped.



4. If we add an alternative, the probability for other alternatives drops.

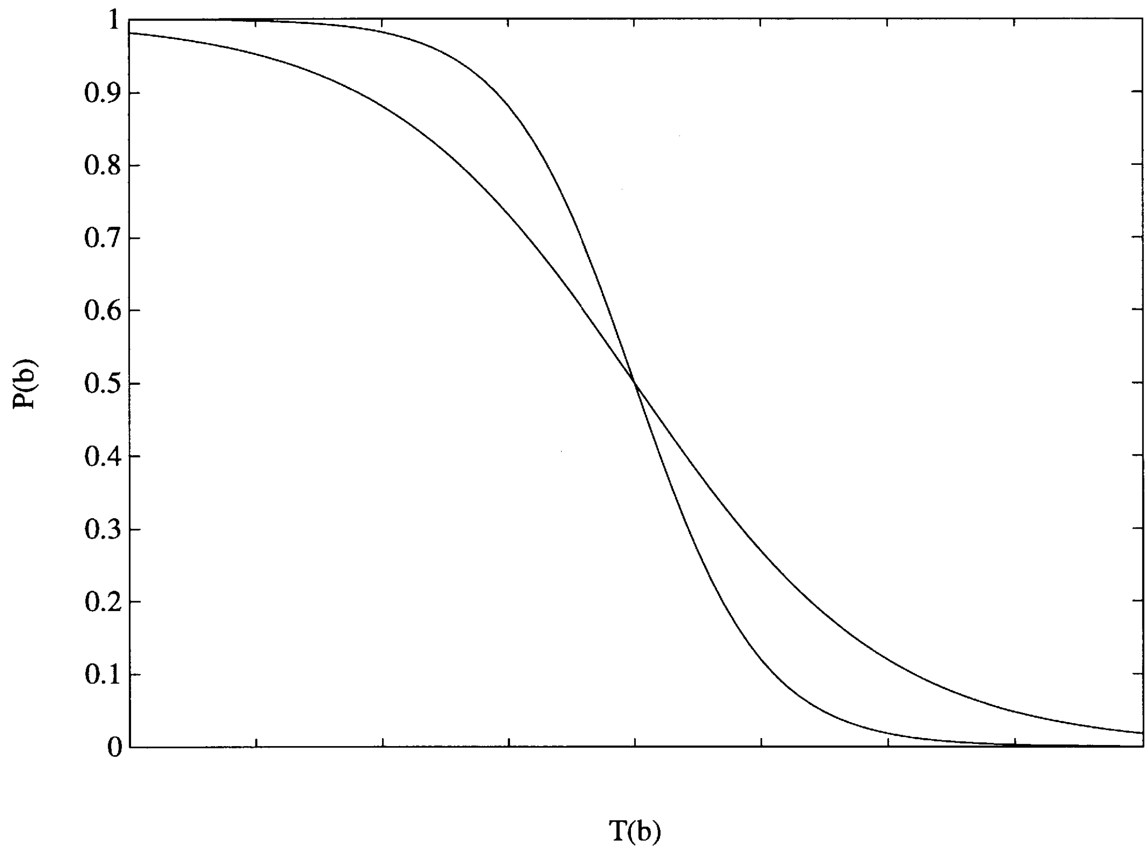
   New mode: rail.

$$P_b = \frac{e^{\alpha C_b + \beta T_b}}{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b} + e^{\alpha C_r + \beta T_r}}$$

$$< \frac{e^{\alpha C_b + \beta T_b}}{e^{\alpha C_c + \beta T_c} + e^{\alpha C_b + \beta T_b}}$$

Parameters

Magnitude of the parameters determines steepness of the S-curve.



Behavioral Motivation

Behavior

U_{in} = utility that person *n* obtains from alternative *i*.

Person chooses alternative with the highest utility.

Econometrics

Decompose: $U_{in} = V_{in} + \varepsilon_{in}$

$V_{in} = V(z_{in}, s_n; \beta)$

= known function of observed data, called ***representative utility***

ε_{in} = unknown component of utility

Then:

$P_{in} = \text{Prob}(U_{in} > U_{jn} \text{ for all } j \in J_n)$

$P_{in} = \text{Prob}(V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn} \text{ for all } j \in J_n)$

$P_{in} = \text{Prob}(\varepsilon_{jn} - \varepsilon_{in} < V_{in} - V_{jn} \text{ for all } j \in J_n)$

Logit

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

If

$$\varepsilon_{in} \sim \text{extreme value}$$

$$\text{Density of } \varepsilon_{in} = e^{-\varepsilon_{in}} e^{-e^{-\varepsilon_{in}}}$$

Then

$$P_{in} = \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}}$$

Example

Commuter's Choice of Mode

Suppose: $V_c = \alpha C_c + \beta T_c = 4$

$$V_b = \alpha C_b + \beta T_b = 3$$

$$U_c = 4 + \varepsilon_c$$

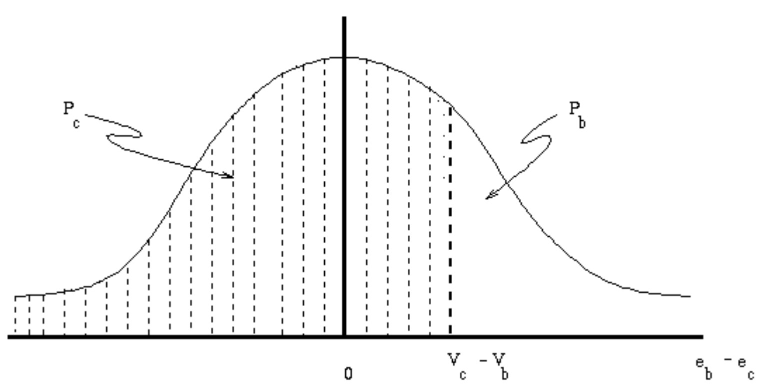
$$U_b = 3 + \varepsilon_b$$

Person chooses car if ε_b is less than ε_c
or if ε_b exceeds ε_c by less than one.

Person chooses bus if ε_b exceeds ε_c by more than one.

That is: $P_c = \text{Prob}(\varepsilon_b < \varepsilon_c + 1) = \text{Prob}(\varepsilon_b - \varepsilon_c < 1)$

More generally: $P_c = \text{Prob}(\varepsilon_b - \varepsilon_c < V_c - V_b)$



Meaning of Parameters

$$U = \alpha C + \beta T + \varepsilon$$

$$dU = \alpha dC + \beta dT = 0$$

$$\text{Value of time} = -\left[\frac{dC}{dT}\right]_{\bar{u}}$$

$$\alpha dC = -\beta dT$$

$$\frac{dC}{dT} = -\frac{\beta}{\alpha}$$

Example: $U = -.24 \cdot C - 1.8 \cdot T$

C in dollars, T in hours

$$\text{Value of time} = \frac{1.8}{.24} = \$7.50/\text{hour}$$

Estimation of Parameters

Sample: $n = 1,\ldots,N$

Denote:

$$d_{in} = \begin{cases} 1 & \text{if person n chose alternative i,} \\ 0 & \text{otherwise .} \end{cases}$$

Probability of chosen alternative for person n:

$$\prod_i P_{in}^{d_{in}}$$

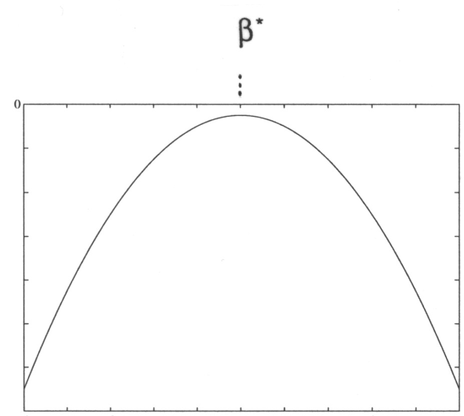
Likelihood function:

$$L = \prod_n \prod_i P_{in}^{d_{in}}$$

Log likelihood function:

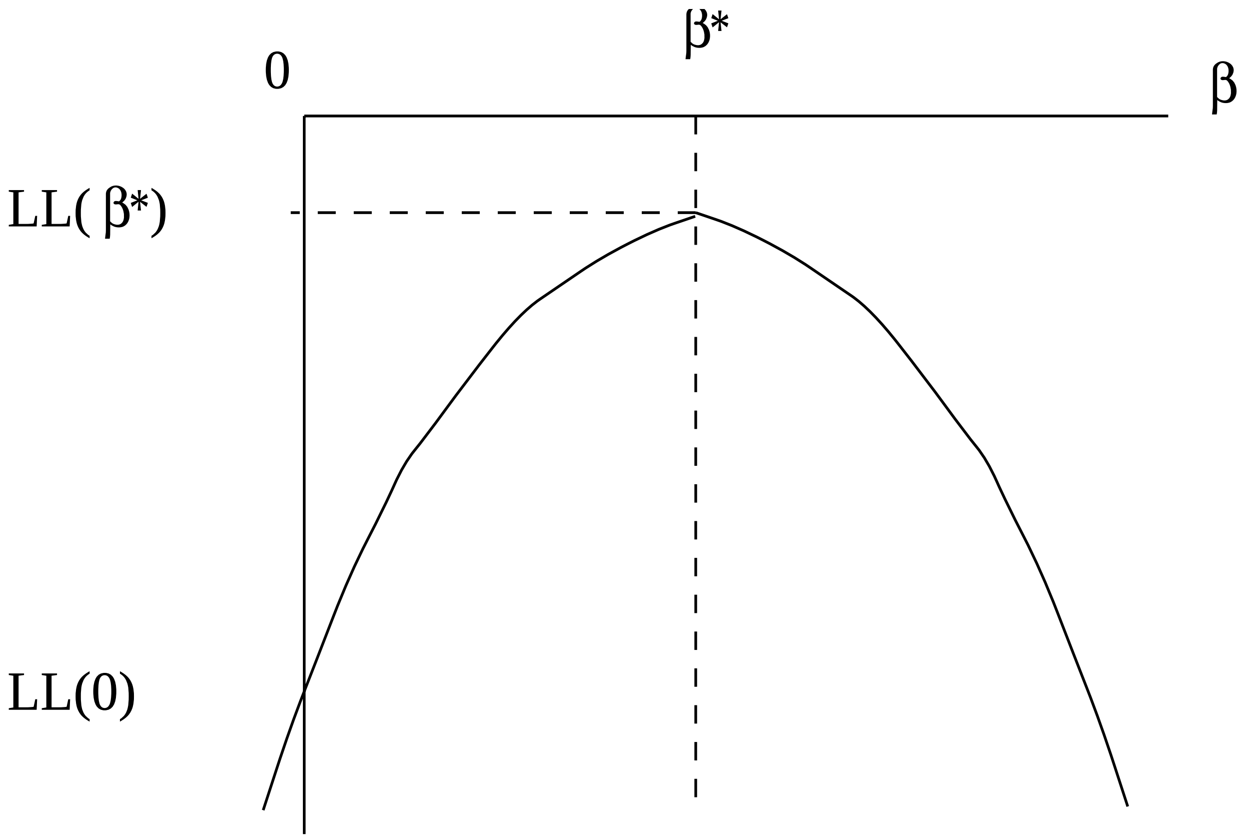
$$LL = \log(L) = \sum_n \sum_i d_{in} \log P_{in}$$

Maximum likelihood estimation:



Goodness of Fit

Likelihood Ratio Index



$$\mu = 1 - \frac{LL(\beta^*)}{LL(0)}$$

If model is perfect, $LL(\beta^*) = 0$ and $\mu = 1$.

If model is no better than chance, $LL(\beta^*) = LL(0)$ and $\mu = 0$.