Topics

<u>Discrete Dependent Variables Models</u>

Multinomial Data – MN Logit & MN Probit

Ordered Data – Ologit & Oprobit

Multivariate Data – MV Probit

<u>Limited Dependent Variables Models</u>

Counted Data – Poisson Regression Model

Truncated Distribution – Truncated Regression Model

Censored Distribution – Tobit Model

By Tatre Jantarakolica

Multinomial Data

Dependent variable has a finite number of possible outcomes.

Individuals can choose among more than two choices.

Individual can choose only one choice.

Multinomial Data

Unordered → IIA valid?

- Yes → MN Logit
- No → MN Probit, Nested logit

Ordered → Ordered logit/probit

where: IIA is Independence of Irrelevant Alternatives (assumption)

Multinomial Data

Multiple alternatives without obvious ordering

Choice of a single alternative out of a number of distinct alternatives.

Example of Multinomial Data:

Which mode of transportation do you use to get to work? – bus, car, bicycle etc.

Example of Ordered Data:

How do you feel today: very well, fairly well, not too well, miserably

Multinomial Logit

Random Utility

A discrete choice underpinning

choice between M alternatives

decision is determined by the utility level Uij, an individual i derives from choosing alternative j

Let: $$U_{ij} = x_{ij}\beta_j + \varepsilon_{ij}$$

where $i=1,\ldots,N$ individuals; $j=0,\ldots,J$ alternatives

The alternative providing the highest level of utility will be chosen.

Multinomial Logit

The probability that alternative j will be chosen is:

$$P(y_i = j) = P(U_{ij} \geq U_{ik} \mid x, \forall k \neq j)$$

$$= P(\varepsilon_{ik} - \varepsilon_{ij} \leq x_{ij}\beta_j - x_{ij}\beta_k \mid x, \forall k \neq j)$$

In order to calculate this probability, the maximum of a number of random variables has to be determined.

In general, this requires solving multidimensional integrals → analytical solutions do not exist.

Multinomial Logit

Exception: If the error terms ε_{ij} are assumed to be independently & identically standard extreme value distributed, then an analytical solution exists.

In this case, similar to binary Logit, it can be shown that the choice probabilities are

$$P(y_i = j) = \frac{\exp\left(x_{ij}\beta_j\right)}{\sum_k \exp\left(x_{ik}\beta_k\right)}$$

Multinomial Logit

By setting up the based case, the estimated results will be compared with the based case. Then:

$$P(y_i = j) = \frac{\exp(x_{ij}\beta_j)}{1+\sum_k \exp(x_{ik}\beta_j)}, k = 1,\ldots,K,\ j = 2,\ldots,J$$

$$P(y_i = 1) = \frac{1}{1+\sum_k \exp(x_{ik}\beta_j)}, k = 1,\ldots,K$$

The special case where $j=1$ yields the binary Logit model.

Multinomial Logit

<u>Example</u>: For the case of three choices

$$P(y_i = 1) = \frac{\exp(x_i\beta_1)}{\exp(x_i\beta_1) + \exp(x_i\beta_2) + \exp(x_i\beta_3)}$$

$$P(y_i = 2) = \frac{\exp(x_i\beta_2)}{\exp(x_i\beta_1) + \exp(x_i\beta_2) + \exp(x_i\beta_3)}$$

$$P(y_i = 3) = \frac{\exp(x_i\beta_3)}{\exp(x_i\beta_1) + \exp(x_i\beta_2) + \exp(x_i\beta_3)}$$

Compare case 2 & 3 with 1 as based case

$$P(y_i = 1) = \frac{1}{1 + \exp(x_i(\beta_2 - \beta_1)) + \exp(x_i(\beta_3 - \beta_1))}$$

$$P(y_i = 2) = \frac{\exp(x_i(\beta_2 - \beta_1))}{1 + \exp(x_i(\beta_2 - \beta_1)) + \exp(x_i(\beta_3 - \beta_1))}$$

$$P(y_i = 3) = \frac{\exp(x_i(\beta_3 - \beta_1))}{1 + \exp(x_i(\beta_2 - \beta_1)) + \exp(x_i(\beta_3 - \beta_1))}$$

Multinomial Logit

Different kinds of independent variables:

- Characteristics that do not vary over alternatives (e.g., socio-demographic characteristics, time effects)
- Characteristics that vary over alternatives (e.g., prices, travel distances etc.)

In the latter case, the multinomial logit is often called "conditional logit".

It requires a different arrangement of the data (one row per alternative for each i)

Multinomial Logit

The model:

$$U_i^j = u_{ij} + \varepsilon_{ij} = x_i'\beta_j + \varepsilon_{ij}$$

β_j measures the relative weights of the characteristics in the derived utility.

$\beta_{jl} - \beta_{hl}$ measures the marginal increase of the utility of alternative j as compared to alternative h when the lth explanatory variable raises by one unit.

Conditional Logit

Another type of model is when aspects of the alternatives are measured for each individual – e.g. interest rates among financial strategies.

$$U_i^j = u_{ij} + \varepsilon_{ij} = x_{ij}'\beta + \varepsilon_{ij}$$

where:

x_{ij} is vector of explanatory variables that apply for individual i and alternative j.

$x_{ij} - x_{ih}$ measures the differences between alternative j and alternative h – e.g. differences in interest rates between strategy j and h.

Multinomial and Conditional Logit Model

Assume that the error terms are independently and identically distributed.

Multinomial Logit:
$$p_{ij} = \frac{e^{x_i'\beta_j}}{\sum_{h=1}^{m} e^{x_i'\beta_h}} = \frac{e^{x_i'\beta_j}}{1+\sum_{h=2}^{m} e^{x_i'\beta_h}}$$

Conditional Logit:
$$p_{ij} = \frac{e^{x_{ij}'\beta}}{\sum_{h=1}^{m} e^{x_{ih}'\beta}}$$

Choice Model and Log-likelihood

For both models, assume maximize utility.

$$p_{ij} = \Pr[y_i = j] = \Pr[u_{ij} + \varepsilon_{ij} > u_{ih} + \varepsilon_{ih} \;\; \textit{for all} \;\; h \neq j]$$

Assume individuals make *independent* choice, the log-likelihood function can be stated as:

$$\log L = \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} \log p_{ij} = \sum_{i=1}^{n} \log p_{iy_i}$$

where:

$y_{ij} = 1$ if $y_i = j$ and $y_{ij} = 0$ otherwise.

Note: This loglikelihood function is globally concave and easy to maximize (McFadden, 1974) → big computational advantage over multinomial probit or nested logit.

Multinomial Logit

Interpretation of coefficients

The coefficients themselves cannot be interpreted easily but the exponential coefficients have an interpretation as the relative risk ratios (RRR)

Let $P(y_i = j) = p_{ij}$

$$\frac{p_{ij}}{p_{i0}} = \exp(x_{ij}\beta_j)$$ "risk ratio"

(for simplicity, only one regressor considered)

Multinomial Logit

The relative risk ratio tells us how the probability of choosing *j* relative to *0* changes if we increase *x* by one unit:

Such that $$\frac{p_{ij}'}{p_{i0}'} = \exp((x_{ij}+1)\beta_j)$$

Relative Risk Ratio (RRR) = $$\exp(\beta_j) = \frac{p_{ij}'}{p_{i0}'} \bigg/ \frac{p_{ij}}{p_{i0}}$$

Multinomial Logit

Interpretation:

Variable x increases (decreases) the probability that alternative j is chosen instead of the baseline alternative if RRR > (<) 1.

Multinomial Logit

Marginal effects of the explanatory variables on the choice probabilities:

Multinomial Logit:

$$\frac{\partial \Pr_{MNL}[y_i = j]}{\partial x_i} = p_{ij}\left(\beta_j - \sum_{h=2}^{m} p_{ih}\beta_h\right)$$

Conditional Logit:

$$\frac{\partial \Pr_{CL}[y_i = j]}{\partial x_{ij}} = p_{ij}\left(1 - p_{ij}\right)\beta$$

$$\frac{\partial \Pr_{CL}[y_i = j]}{\partial x_{ih}} = -p_{ij}\, p_{ih}\beta \quad for \quad h \neq j$$

Multinomial Logit

Elasticities:

$$\frac{\partial P_{ij}}{\partial x_{ik}} \frac{x_{ik}}{P_{ij}} = x_{ik}\left(\beta_{jk} - \sum_{m=1}^{J-1} P_{im}\beta_{mk}\right), \quad j = 1, \ldots, J-1$$

Represent relative change of p_{ij} if x increases by 1 percent

Multinomial Logit

Independence of Irrelevant Alternatives (IIA) is the important assumption of Multinomial Logit Model

It implies that the decision between two alternatives is independent from the existence of more alternatives

Ratio of the choice probabilities between two alternatives j and k is independent from any other alternative:

$$\frac{p_{ij}}{p_{ik}} = \exp\left[x_{ij} (\beta_j - \beta_k) \right]$$

Multinomial Logit

Problem: This assumption is invalid in many situations.

Example: Red bus – Blue bus – Problem.

Initial situation: An individual chooses to

- walk with probability 2/3
- taking the bus (w/o Blue bus) is 1/3

Then, the probability ratio = 2:1

Multinomial Logit

Example: Red bus – Blue bus – Problem.

With the blue buses:

It is rational to believe that the probability of walking will not change.

If the number of red buses = number of blue buses: Person walks with P=4/6

Person takes a red bus with P=1/6

Person takes a blue bus with P=1/6

New probability ratio: 4:1

Not possible according to IIA (IIA violation).

Multinomial Logit

The following probabilities result from the IIA-assumption:

P(by foot)=2/4

P(red bus)=1/4

P(blue bus)=1/4, such that $\frac{\text{P(walk)}}{\text{P(red bus)}} = \frac{2}{1}$

Problem:

Probability of walking decreases from 2/3 to 2/4 due to the introduction of blue buses

This is not plausible.

Multinomial Logit

Reason of IIA property assumes that error terms are independently distributed over all alternatives.

The IIA property causes no problems if all alternatives considered differ in almost the same way.

e.g., probability of taking a red bus is highly correlated with the probability of taking a blue bus

This is just the substitution patterns.

Multinomial Logit

Test of IIA:

Hausman Test:

H_0: IIA is valid ('odds ratios' are independent of additional alternatives)

Procedure: "omit" a category – Do the estimated coefficients change significantly?

If Yes → reject H_0 → Multinomial Logit should not be applied.

The model should be Nested Logit or Multinomial Probit instead.

Multinomial and Conditional Probit Model

Assume joint normal distribution for the model.

$$\begin{pmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{im} \end{pmatrix} \sim NID(0,V)$$

If $u_{ij} = x_i \beta_j$, the model is multinomial probit model.

If $u_{ij} = x_{ij} \beta$, the model is conditional multinomial probit model.

Ordered Response Model

In case that choices are ordered, ordered probit or logit model can be applied.

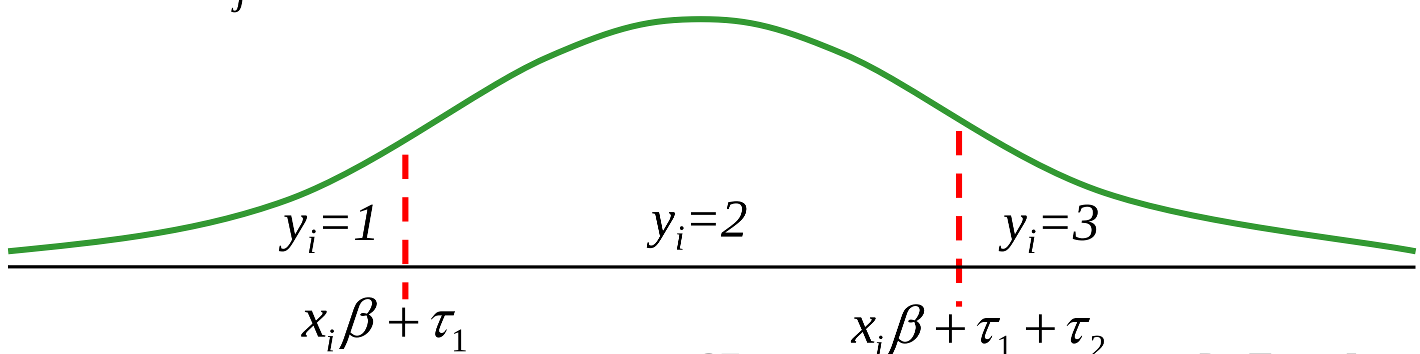
$$y_i^* = x_i\beta + \varepsilon_i, \quad E(\varepsilon_i) = 0$$

$$y_i = 1 \quad if \; -\infty < y_i^* \le \tau_1,$$

$$y_i = j \quad if \sum_{k=1}^{j-1} \tau_k < y_i^* \le \sum_{k=1}^{j} \tau_k, \; j = 2,\ldots,m-1,$$

$$y_i = m \quad if \sum_{k=1}^{m-1} \tau_k < y_i^* \le \infty$$

where τ_j = Threshold value (cut), $j=1,2,\ldots, m$



By Tatre Jantarakolica