A Unimodal Ordered Logit model for ranked choices

Melvin Wong[1] José Ángel Martín-Baos[2] Michel Bierlaire[1]

[1]Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
École Polytechnique Fédérale de Lausanne

[2]Department of Mathematics
University of Castilla-La Mancha, Spain

12 September 2021

Outline

- Introduction
- Background
- Methodology: Unimodal logit
- Case study: Crash severity model
- Conclusion

Introduction

- Ordinal scale responses capture qualitative user feedback
- Responses have inherent correlation between alternatives [Small, 1987]

Examples

PT satisfaction, driver star-rating (ride-hailing), crash severity...

- [Krueger et al., 2019, Tirachini and del Río, 2019, Fu, 2020, Loa and Habib, 2021]

Background

[McCullagh, 1980]

Proportional odds model

- Contiguous intervals on a continuous scale
- Points of division assumed to be unknown

↓

[Small, 1987]

Ordered logit, Generalized ordered logit

- Define a latent variable (y^*) that varies across the contiguous intervals
- $y^* \leftarrow$ exogenous features of the response
- $y^* = \sum_m \beta_m X_m$
- Choice prob. = probability of lying in any of the intervals

Modelling non-ordered choices

Assume that there are J alternatives ($i = 1, ..., J$)

- Denote $y_{ni} = 1$ if individual n is ranked in i and $y_{ni} = 0$ otherwise
- $n = 1, ..., N,\ U_{n1}, ..., U_{nJ},\ U_{ni} \geq \max\{U_{n1}, ..., U_{nJ}\}$
- $U_{ni} = V_{ni} + \varepsilon_{ni}, \quad \varepsilon_{ni} \sim \text{Gumbel}(0, 1)$ i.i.d.

Multinomial logit model

$$P(y_{ni} = 1) = \frac{\exp(V_{ni})}{\sum_{j=1}^{J} \exp(V_{nj})}$$

For choices with natural ordering, i.i.d. assumption does not hold

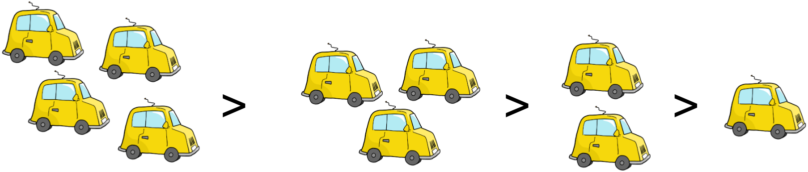
- Standard MNL model is not suitable in this context of ranked choices

Modelling ordered choices

Example

Vehicle ownership model [Sheffi, 1979]

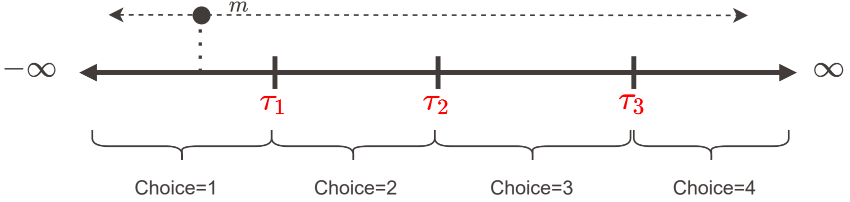
- j is the number of household vehicles (j =1, 2, 3,...)
- n would prefer j over $j-1$, $j-1$ over $j-2$, and so on..
- Correlation between choices not captured (MNL)



Modelling ordered choices

Ordered logit

$$y_n^* = \sum_m \beta_m X_{mn} + \varepsilon_n$$



Estimating thresholds

$\tau_1 = 0$, $\tau_2 = \tau_1 + \Delta_2$, $\tau_3 = \tau_2 + \Delta_3$

Modelling ordered choices

Ordered logit

- The difference between thresholds (e.g. between τ_2 and τ_3) are assumed to be the same for all respondents
- Parameters β_m are constant across all respondents
- Typically set threshold $\tau_1 = 0$ for model identification

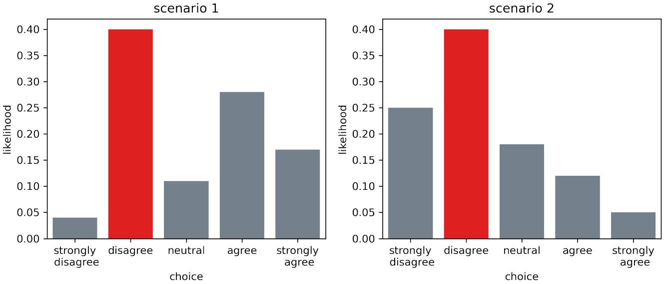
Generalized ordered logit [Eluru et al., 2008]

- latent variable combines alt. specific and generic parameters
  $y_n^* = \sum_m \beta_m X_{mn} + \sum_m \beta_{im} X_{mn} + \varepsilon_n$
- Thresholds are functions of exogenous variables:
  $\tau_i = \tau_{i-1} + \exp(\sum_m \delta_{im} Z_{imn})$

Other models for ordered choices

- Generalized Extreme Value (GEV) [McFadden, 1977]
- Ordered GEV [Small, 1987]
- Dogit model [Gaundry and Dagenais, 1979]
- Dogit OGEV model [Fry and Harris, 2005]

A different approach for ordered choices



Maximum likelihood estimation

$$\ln(P(y_{\text{scenario }1} = 2)) = \ln(P(y_{\text{scenario }2} = 2))$$

- Both result in identical max likelihood, but probability mass function (pmf) is different

Unimodality in ordered choices

Properties

- Unimodality: A single highest value
- Specifically, the *a posteriori* choice probabilities are unimodal

Natural ordering of choices is captured in the model if there exist an integer $c \in J$ such that:

- $p(y_{ni}|X) \geq p(y_{ni+1}|X)$, for all $i \geq c$ and,
- $p(y_{ni-1}|X) \leq p(y_{ni}|X)$, for all $i \leq c$

Unimodality in ordered choices

Poisson pmf

The probability of i occurrences of an event in a set of N observations is defined as:

$$P(i) = \frac{\lambda^i \exp(-\lambda)}{i!}, \text{ for } \quad i = 0, 1, 2, ..$$

Unimodal logit

Applying a unimodal constraint in the utility function:

$$\begin{aligned} U_{in} &= V_{in} + \ln(P(i)) + \varepsilon_{in} \\ &= V_{in} + \ln\left(\frac{\lambda^i \exp(-\lambda)}{i!}\right) + \varepsilon_{in} \\ &= V_{in} + \underbrace{i\ln(\lambda) - \lambda - \ln(i!)}_{\text{error component } f(\lambda,i)} + \varepsilon_{in} \end{aligned}$$

ec: capture correlations among utilities of alternatives

Conditions

- λ is positive
- $\lambda = f(y_n^*) = \ln(1 + \exp(y_n^*))$

Unimodal logit

Expressed as a MNL choice probability:

$$P(y_{ni} = 1) = \frac{\exp(\mu\Phi_{in})}{\sum_{j=1}^{J} \exp(\mu\Phi_{jn})}$$

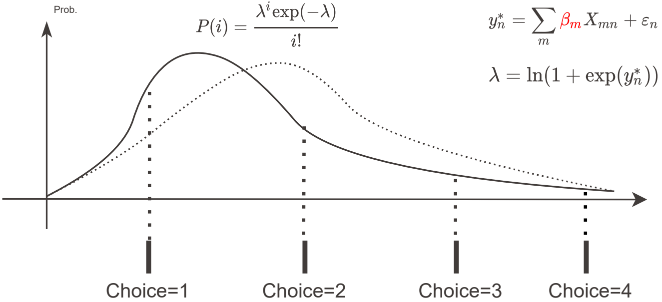
$$\Phi_{in} = V_{in} + i\ln(\lambda) - \lambda - \ln(i!) + \beta_{i0}$$

Behavioural interpretation

Utilities of alternatives are corrected for proximity from the selected choice i

Unimodal logit

Unimodal logit



Zero-truncated Poission (ZTP) pmf

When a choice set has a "zero" option

- Example: Number of items in a shopping cart include a "no purchase" option

A ZTP Unimodal logit has the following pmf:

$$P(i|i>0) = \frac{\lambda^i \exp(-\lambda)}{i!(1-\exp(-\lambda))}, \text{ for } i = 1, 2, 3, \ldots$$

$$U_{in} = V_{in} + i\ln(\lambda) - \lambda - \ln(i!) - \ln(1-\exp(-\lambda)) + \varepsilon_{in}$$

Crash severity model

[City of Tempe, 2018]

Open dataset: High Severity Traffic Crash Data Report

- 39,793 records (2012–2019)
- Five severity levels
  1: No injury, 2: possible injury, 3: minor injury, 4: major injury, 5: fatal
- 28 crash and environmental features used (after data cleaning)

Models

Estimation using Biogeme [Bierlaire, 2020]

- Ordered logit
- Unimodal logit
- Zero truncated unimodal logit

Crash severity model

Model Evaluation

Goodness-of-fit

- Pseudo R-squared measure (ρ^2)

  $$\rho^2 = 1 - \frac{\ln LL(\hat{\beta})}{\ln LL(\bar{\beta})}$$

- Bayesian Information Criterion (BIC)

  $$BIC = -2LL(\beta) + M \ln(Q)$$

Out-of-sample accuracy

- Discrete classification accuracy
- Geometric mean probability of correct assignment (GMPCA) [Hillel, 2019]
- Quadratic Weighted Kappa (QWK) [Cohen, 1968]

Model results

Abridged results (1)

Variables	Ordered Logit values	Ordered Logit rob_tTest	Unimodal values	Unimodal rob_tTest	Zero-trunc Unimodal values	Zero-trunc Unimodal rob_tTest
age	-0.008	-10.184	-0.015	-22.834	-0.017	-18.351
alcohol	0.384	5.02	0.379	4.625	0.524	4.918
cause_distraction	**0.08**	**1.013**	**-0.287**	**-4.749**	-0.249	**-2.83**
cause_speeding	-0.027	**-0.543**	-0.271	**-6.78**	-0.28	**-4.824**
cause_turn	-0.153	-1.832	-0.355	-6.05	-0.411	-4.611
cause_yield	-0.108	-2.038	-0.341	-8.166	-0.4	-6.784
type_cyclist	1.46	**17.722**	0.619	**5.289**	0.804	**6.265**
type_driverless	-0.52	**-1.465**	-1.478	**-7.631**	-1.744	**-5.267**
type_pedestrian	1.596	7.657	3.838	8.122	3.066	6.065

Model results

Abridged results (2)

	Ordered Logit		Unimodal		Zero-trunc Unimodal	
Variables	values	rob_tTest	values	rob_tTest	values	rob_tTest
ASC_noinjury (1)			ref.		ref.	
ASC_possinjury (2)			3.673	93.809	2.452	68.248
ASC_nonincap (3)			4.117	104.798	3.446	97.33
ASC_incap (4)			2.449	43.533	1.907	35.544
ASC_fatal (5)			0.788	7.193	0.319	2.889
tau1	0.0	0.0				
delta2	2.611	68.111				
delta3	3.31	39.596				
delta4	2.303	14.98				

Model results

	Ordered Logit	Unimodal	Zero-trunc Unimodal
Log likelihood	-17148.44	**-13471.31**	-16731.04
BIC	34628.6	**27274.4**	33793.8
ρ^2	0.665	**0.737**	0.673
Optimization time	0:01:02.27	0:06:26.2	0:07:40.4
Discrete Class. Acc.	0.839	0.842	0.826
GMPCA	0.581	**0.653**	0.59
QWK	0.758	**0.805**	0.787
20% out-of-sample data used			

Conclusion

We introduce a new form of choice model for ordered choices

- Unimodal constraint on the *a posteriori* distribution
- Similar β interpretations as Ordered logit

Case study

- Able to capture the influence of relevant crash severity characteristics: driving speed, distracted driving and driverless vehicles
- Exhibit better model fit and forecasting accuracy

Future work

- Negative binomial distribution
- Combination with other error correction functions

Thank you for your attention

Estimated models, cleaned data and data analysis are available at:
`https://github.com/mwong009/unimodal-logit`

References I

Michel Bierlaire. A short introduction to pandasbiogeme. Technical Report TRANSP-OR 200605, Transport and Mobility Laboratory, ENAC, EPFL, 2020.

City of Tempe. Crash data report (detail), 2018. URL `https://open.tempe.gov/datasets/1-08-crash-data-report-detail`.

Jacob Cohen. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4):213, 1968.

Naveen Eluru, Chandra R Bhat, and David A Hensher. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis & Prevention*, 40(3):1033–1054, 2008.

Tim RL Fry and Mark N Harris. The dogit ordered generalized extreme value model. *Australian & New Zealand Journal of Statistics*, 47(4):531–542, 2005.

Xue-mei Fu. Does heavy ict usage contribute to the adoption of ride-hailing app? *Travel Behaviour and Society*, 21:101–108, 2020.

Marc JI Gaundry and Marcel G Dagenais. The dogit model. *Transportation Research Part B: Methodological*, 13(2):105–111, 1979.

Tim Hillel. *Understanding Travel Mode Choice: A New Approach for City Scale Simulation*. Thesis, University of Cambridge, 2019.

References II

Rico Krueger, Taha H Rashidi, and Joshua Auld. Preferences for travel-based multitasking: Evidence from a survey among public transit users in the chicago metropolitan area. *Transportation Research Part F: Traffic Psychology and Behaviour*, 65:334–343, 2019.

Patrick Loa and Khandker Nurul Habib. Examining the influence of attitudinal factors on the use of ride-hailing services in toronto. *Transportation Research Part A: Policy and Practice*, 146:13–28, 2021.

Peter McCullagh. Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(2):109–127, 1980.

D McFadden. A closed-form multinomial choice model without the independence from irrelevant alternatives restrictions (working paper no. 7703). *Berkeley. CA: University of California, Urban Travel Demand Forecasting Project, Institute of Transportation Studies*, 1977.

Yosef Sheffi. Estimating choice probabilities among nested alternatives. *Transportation Research Part B: Methodological*, 13(3):189–205, 1979.

Kenneth A Small. A discrete choice model for ordered alternatives. *Econometrica: Journal of the Econometric Society*, pages 409–424, 1987.

Alejandro Tirachini and Mariana del Río. Ride-hailing in santiago de chile: Users' characterisation and effects on travel behaviour. *Transport Policy*, 82:46–57, 2019.