ISSN 2348 – 8034
Impact Factor- 4.022

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

INVESTIGATION AND COMPARISON OF ROUTE CHOICE MODELS

Inchul Yang[1] & Woo Hoon Jeon*[2]
[1]Senior Researcher, Korea Institute of Civil Engineering and Building Technology, Korea
[*2]Senior Researcher, Korea Institute of Civil Engineering and Building Technology, Korea

ABSTRACT
In this study, MNL and C-logit are investigated as GEV type model, and MNP as non-GEV type model. Three models are tested in a very simple network to compare their performance and usefulness in practice. The layout of the paper is as follows. Section 2 provides a theoretical aspect of MNL, C-logit and MNP. Section 3 shows implementation of the models in a simple network and comparison of the results. Section 4 provides conclusion.

***Keywords**:* Route Choice Model, Transportation Road Network, C-Logit, Multinomial Logit, Multinomial Probit

I. INTRODUCTION

The route choice problems have been studied for decades to investigate drivers' behavior in transportation road networks. The endeavors to alleviate traffic congestion have been enriched researches on future traffic prediction in which the route choice behavior plays an important role.

The discrete choice theory is widely adopted to develop the route choice models including multinomial logit (MNL), C-logit, path-size logit, nested logit, paired combinatorial logit, cross-nested logit, generalized nested logit, multinomial probit (MNP), heteroscedastic extreme value logit (HEV) and error components logit (or mixed logit). The former seven models are categorized as GEV-type model, and the latter as non-GEV-type model.

In this study, MNL and C-logit are investigated as GEV type model, and MNP as non-GEV type model. Three models are tested in a very simple network to compare their performance and usefulness in practice. The layout of the paper is as follows. Section 2 provides a theoretical aspect of MNL, C-logit and MNP. Section 3 shows implementation of the models in a simple network and comparison of the results. Section 4 provides conclusion.

II. LITERATURE REVIEW

1. Multinomial Logit (MNL)

MNL is one of the widely employed models to represent urban drivers' route choice behavior. In a choice among path set C_n given measured impedances $V_i\ \forall i \in C_n$, the probability of choosing path i is given by:

$$P_i = \frac{\exp(-\beta_0 V_i)}{\sum_{j=1}^{J} \exp(-\beta_0 V_j)}$$

where:
J : Number of paths;
β_0 : Dispersion parameter to be calibrated.

The parameter, β_0, can be estimated by MLE with the following likelihood function:

[Jeon, 4(10): October 2017] ISSN 2348 – 8034
DOI- 10.5281/zenodo.1009312 Impact Factor- 4.022

$$L = \prod_{n=1}^{N}\prod_{i=1}^{J} P_{ni}^{1\{y_{ni}=i\}} = \prod_{n=1}^{N}\prod_{i=1}^{J}\left(\frac{\exp(-\beta_0 V_{ni})}{\sum_{j=1}^{J}\exp(-\beta_0 V_{nj})}\right)^{1\{y_{ni}=i\}}$$

where:

P_{ni} : Probability of person n choosing path i;

N : Number of persons;

J : Number of paths available for person n.

$1\{y_{ni}=i\}$: Indicator function of the event $y_{ni}=i$, which takes 1 if the event is true and 0 otherwise

2. C-Logit

The C-logit, developed by Cascetta et al. (1996), is one of the path-enumeration based route choice models derived from McFadden's generalized extreme value (GEV) theory. It accommodates similarities between overlapping paths through an additional attributes (the commonality factor) in the utility function. The commonality factor reduces the utility of overlapping paths and increases the utility of independent paths.

The model formulation is as follows:

$$P(i\,|\,C_n) = \frac{\exp(\beta_0(V_{in} - CF_{in}))}{\sum_{j=1}^{J}\exp(\beta_0(V_{jn} - CF_{jn}))}$$

where:

V_{in} : Utility of path i for person n;

C_n : Path set for person n;

CF_{in} : Commonality factor of path i for person n;

β_0 : Parameter to be calibrated.

The commonality factor of path i, is a measure of the degree of similarity of path i with other paths in an OD pair. The commonality factor can be specified in different ways. Cascetta et al. (1996) proposed the following specification:

$$CF_{in} = \beta_1 \ln \sum_{j\in C_n}\left(\frac{L_{ij}}{\sqrt{L_i L_j}}\right)^{\gamma}$$

where:

L_{ij} : Length of links common to path i and path j;

L_i, L_j : Overall path lengths of path i and path j, respectively;

β_1 and γ : Parameters to be calibrated.

Recently, Cascetta (2001) presented the following alternative formulation for the commonality factor:

$$CF_{in} = \beta_0 \ln\left[1 + \sum_{\substack{j \in C_n \\ j \neq i}} \left(\frac{L_{ij}}{\sqrt{L_i L_j}}\right)\left(\frac{L_i - L_{ij}}{L_j - L_{ij}}\right)\right]$$

3. Multinomial Probit (MNP)

According to MNP the choice probability in general is given by:

$$P_{ni} = \Pr ob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}, \forall j \neq i)$$

where:

$\langle \varepsilon_n \rangle \sim MVN(0, \Sigma)$: Multivariate normal vector

Due to an absence of a closed form of MNP, GHK simulator is adopted to yield the probability of choosing paths.

4. GHK Simulator

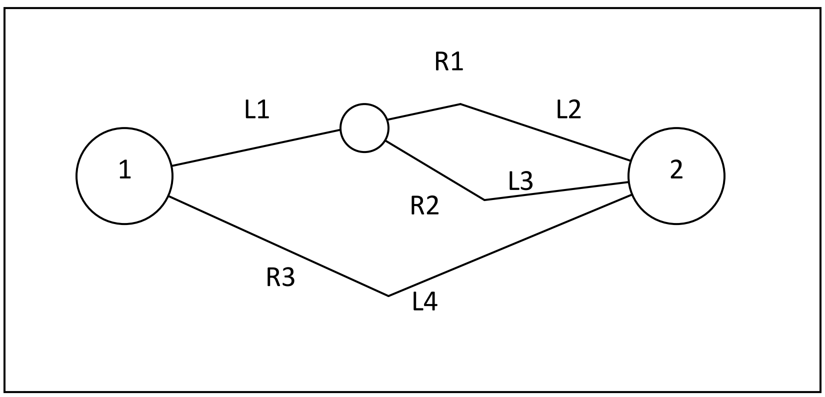
The GHK simulator, after Geweke, Hajivassiliou and Keane, is most widely used probit simulator. It is said that the GHK simulator is the most accurate and useful among numerous probit simulator.

The GHK employs utility differences. The simulation of probability Pni starts by subtracting the utility of alternative i from each other alternative's utility. The fact that the utility of a different alternative is subtracted depending on which probability is being simulated is critical to the implementation of the procedure (Train, 2003).
The details can be found in Train (2003).

III. EMPIRICAL COMPARISON

1. Network simulation

A test network consists of four links (L1, L2, L3 and L4), with three route alternatives (R1, R2 and R3) for traveling between origin O and destination D. These were specified as follows (with mean impedances given in brackets):



R1 = L1(2) + L2(3)
R2 = L1(2) + L3(3.5)
R3 = L4(6)

Data sets made up of 10000 simulated route choices on the network were constructed as follows. The utility function for route Rn was specified:

[Jeon, 4(10): October 2017] ISSN 2348 – 8034
DOI- 10.5281/zenodo.1009312 Impact Factor- 4.022

$$U_n = \sum_m \beta I_m$$

where:

β : parameter associated with impedance (set equal to -1)

I_m : impedance of link m given by:

$$I_m = \mu_m (1 + X_m \theta)(1 + Y_m \rho)$$

where:

μ_m : fixed mean impedance of link m;

X_m : 'known' sampled from the standard Normal distribution;

θ and ρ : parameters (set equal to 0.3);

Y_m : 'unknown' sampled from the standard Normal distribution;

For each observation, the route with minimum 'known' impedance was indicated as chosen.

2. Model estimation

For the data set, MNL, C-logit and MNP models are estimated. MNL and C-logit are estimated by maximum likelihood estimation and MNP by the maximum simulated likelihood method. In MNP estimation, 100 pseudo-random draws are used to simulate the likelihood function.

Train's procedure (2003) is employed to address some issues when using the GHK simulator in maximum likelihood estimation. The issues are summarized as follows:

1) Different utility differences must be taken for decision makers who chose different alternatives.
2) Denote $\tilde{\Omega}_i$ as covariance matrix of the error differences against i. It must be assured that the parameters in $\tilde{\Omega}_i$ are consistent with those in $\tilde{\Omega}_j$, in the sense that they both are derived from a common Ω.
3) It must be assured that the parameters that are estimated by maximum likelihood imply covariance matrices $\tilde{\Omega}_j \ \forall j$ that are positive definite.
4) It must be assured that the model is normalized for scale and level of utility, so that the parameters are identified.

To assure that the model is identified, in Ω (3x3 matrix), I assume that σ_{22} and σ_{12} are variables, and others fixed such as $\sigma_{11}=\sigma_{33}=1$ and $\sigma_{13}=\sigma_{23}=0$. It's reasonable because only path 1 and path 2 have a common link, L1.

From the Ω, we can get $\tilde{\Omega}_1$ by using M_1 matrix as follows:

[Jeon, 4(10): October 2017] ISSN 2348 – 8034
DOI- 10.5281/zenodo.1009312 Impact Factor- 4.022

$$\tilde{\Omega}_1 = M_1 \Omega M_1^{'} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \sigma_{12} & 0 \\ \sigma_{12} & \sigma_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-2\sigma_{12}+\sigma_{22} & 1-\sigma_{12} \\ 1-\sigma_{12} & 2 \end{pmatrix}$$

Note that M_i is the (J-1) identity matrix with an extra column of -1's added as the ith column. The extra column makes the matrix have size J-1 by J.

To assure that the covariance matrix is positive definite, the model is parameterized in terms of the Choleski factor of $\tilde{\Omega}_1$. That is, a lower triangular matrix that is (J-1)×(J-1) and whose top-left element is 1 is the start point of MLE.

$$L_1 = \begin{pmatrix} 1 & 0 \\ c_{21} & c_{22} \end{pmatrix}$$

where:
c_{21} *and* c_{22} : Parameters to be estimated.

The matrix Ω for the J non-differenced errors is created from L_1. With a matrix, J×J Choleski factor for Ω by adding a row of zeros at the top of L_1 and a column of zeros at the left, the Ω is calculated as $LL^{'}$. With this Ω, $\tilde{\Omega}_j$ for any j can be derived (Train, 2003).

3. Results

Table 1 shows the estimation results. In terms of goodness of fit (final log-likelihood), MNP is the best model. C-logit provides the next best fit followed by MNL. These results are reasonable because MNP and C-logit take into account of common paths while MNL doesn't. Unlike GEV-type model, MNP can consider correlation between paths, which can explain drivers' route choice behavior better.

Table 1. Estimation Results

	MNL	C-logit	MNP
β 0	0.757	0.764	0.447
β 1	-	-0.090	-
σ 12	-	-	0.027
σ 22	-	-	0.302
Final LL	-7644.005	-7643.497	-7585.825
Elapsed Time (millisecond)	989	20028	7952948

In terms of performance, MNL is the best as we can easily expect. MNL has the simplest model, so that it doesn't require a bunch of calculations. Meanwhile, MNP takes about two hours. It's because 100 simulations are required to produce the probability of each data.

[Jeon, 4(10): October 2017] ISSN 2348 – 8034
DOI- 10.5281/zenodo.1009312 Impact Factor- 4.022

IV. CONCLUSION

In this research, MNL and C-logit were investigated as GEV-type model, and MNP as non-GEV-type model. A simple network was used to implement three models, but it is enough to represent the real world road network having common links between more than two paths. MLE was employed to estimate MNL and C-logit, and maximum simulated likelihood estimation method for MNP. The results showed that MNP is the best model to explain drivers' route choice behavior in terms of the goodness of fit (Log-likelihood) followed by C-logit and MNL. It would be interesting to investigate other models such as mixed logit and HEV in the future

REFERENCES

1. *BatleyR., T. Fowkes, D. Watling, G. Whelan, A. Daly, and E. Hato (2001). Models foranalysing route choice. Paper presented at Universities' Transport Studies Group,*
2. *University of Oxford, January 2001.*
3. *Ben-Akiva M., M. S. Ramming and S. Bekhor (2001). Route Choice Models, Presented at theWorkshopon "Human Behaviour and Traffic Networks" , Bonn,December 2001.*
4. *Ben-Akiva M. and M. Bierlaire (1999). Discrete choice methods and their application to short term travel decisions. Handbook of transportation science, R.W. Hall.*
5. *Cascetta E., A. Nuzzolo, F. Russo and A. Vitetta (1996). A Modified Logit Route Choice Model Overcoming Path Overalpping Problems: Specification and Some Calibration Results for Interurban Networks. Transportation and Traffic Theory. Proceedings from the Thirteenth International Symposium on Transportation and Traffic Theory, Lyon, France. J.B. Lesort, ed. Pergamon.*
6. *Cascetta E. and A. Papola (1998). Implicit Availability/Perception Logit Models for Route Choice in Transportation Networks. Paper presented at the 8th World Conference on Transport Research, Antwerp.*
7. *Cascetta E. (2001). Transportation Systems Engineering: Theory and Methods. Applied Optimization Series, 49. Norwell, MA: Kluwer Academic Publishers.*
8. *Train K. (2003). Discrete Choice Methods with Simulation. Cambridge*