

Application of Improved NL Model Based on Dummy Variables in Traffic Mode Selection

Xu He, Wenyong Li

School of Architecture and Transportation Engineering, Guilin University of Electronic Science and Technology, Guilin Guangxi
Email: 798540563@qq.com

Received: May 8th, 2020; accepted: May 22nd, 2020; published: May 29th, 2020

Abstract

Based on the in-depth analysis of the random utility theory and the improvement of the original NL model, an improved NL model based on dummy variables is proposed. The model can select and classify the index of utility selection function again. The property of not quantitative indicators can be injected in the form of dummy variable assignment to the utility equation. This way can make variable regression model to explain widely, thus expanding quantity with good quality, in the face of the description of the realistic problems to have the joint effect that one is worth more.

Keywords

NL Model, Dummy Variables, Traffic Mode Selection

基于哑变量改进的NL模型在交通方式选择中的应用

何 旭, 李文勇

桂林电子科技大学, 建筑与交通工程学院, 广西 桂林
Email: 798540563@qq.com

收稿日期: 2020年5月8日; 录用日期: 2020年5月22日; 发布日期: 2020年5月29日

摘 要

论文通过对随机效用理论进行深度剖析,并在原有NL模型的基础上改进,提出了一种基于哑变量改进的

NL模型。该模型可重新对效用选择函数的指标进行选取和分类, 对于不可量化的性质指标则可以以哑变量赋值的方式注入到效用方程中, 该方式能够使回归模型变量解释范围更广, 进而达到扩量保质的作用, 在面对现实问题的描述时能够起到以一抵多的联合效果。

关键词

NL模型, 哑变量, 出行方式选择

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着我国城镇一体化建设已从慢车道建设步入了快车道, 城市机动车的保有量呈高速几何趋势增长, 交通运输的供给不平衡问题逐渐凸显居[1]。民出行前往往往会结合自身的实际情况来选择不同的交通方式出行, 不同的个体在出行前对于交通方式的选择会带来不同的效用意义[2]。在交通拥堵日益严峻的今天, 通过站在出行者的视角下来解释人们在出行前的选择行为十分必要[3]。

交通出行不仅在时间分布上具有较为明显的周期性, 而且个体选择的差异性往往又充满着离散性[4]。因此在基于简单出行链视角下来引入非集计 NL 模型, 在分类分层的情况下研究居民关注的决策指标对其选择行为所衍生的效用有着重要的研究意义。

2. 基础理论(RUT)

随机效用理论在个体出行的视角下, 通过对出行者所选择的交通方式来综合评判总效用[5]。由于人在进行选择时并非是完全理性的, 在决策者心里必然会存在一定的经验主义。这就使得总的综合效用由两部分组成, 其一是与生活息息相关的可量化的具体效用, 其二是由经验或出行习惯所带来的内在难以描述灰色效用[6]。具体表达式如下所示:

$$U_j = V_j + \varepsilon_j \quad (1)$$

$$V_j = \delta_0 + \delta_1 x_1 + \delta_2 x_2^2 + \dots + \delta_n x_n^n \quad (2)$$

式中: U_j 表示出行者选择第 j 种交通方式出行的总效用; V_j 表示为出行者选择第 j 种交通方式出行的显性效用; x_j 表示为指标变量; ε_j 则表示灰色隐形效用。

3. 建立模型

决策肢的几何结构决定了 NL 模型选择概率的体现形式[7]。因此, 在运用 Nested Logit 模型分析前应将 SP 调查到的数据进行分类筛查, 即: 把出行特征类似的交通工具归为一类。结合本次调查本文将采用两层结构, 其中上层为机动车、非机动车, 下层为小汽车、公交车、电动车。结合上述结构, 建立 NL 模型决策肢方案。

在诸多离散研究中 MNL 模型虽然能够解释决策树中各个方案的选择情况, 但其要求各个备选项之间拥有较强的不相关性, 而这也就是 MNL 所带来的 IIA 局限性[8]。Nested Logit 模型作为离散选择模型中的最经典之一, 其树形结构具有较强的柔性, 弥补了 MNL 在分层分类上各选项间的不相关缺陷[9]。

由于 NL 模型是在 MNL 的基础上改进而来, 故 Nested Logit 也同样服从相同的 Gumbel 假设分布, 各个综合效用函数体现形式仍然具有线性表达特点[10]。决策选择肢示意图见下图 1 所示。

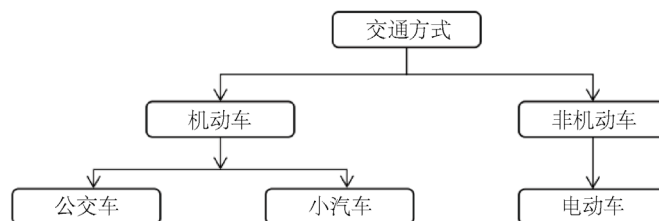


Figure 1. Schematic diagram of transport mode selection

图 1. 交通方式决策示意图

$$\ln \left(\frac{P_j}{1 - P_j} \right) = \alpha + \beta_1 (V_1 + \varepsilon_1) + \beta_2 (V_2 + \varepsilon_2) + \beta_3 (V_3 + \varepsilon_3) + \cdots + \beta_w (V_w + \varepsilon_w) \quad (3)$$

式中: P_j 为居民选择某种交通方式出行的概率; α 为常数项; $\beta_{n=1,2,3,\dots,w}$ 为效用函数系数; ε_w 为隐性效用函数。

$$V_{w=\text{bus,car,fei}} = a_0 + a_1 m_{\text{mean } w} + a_2 d_{\text{mean } w} + a_3 \ln + a_4 sd_1 + a_5 sd_2 \quad (4)$$

式中: $V_{w=\text{bus,car,fei}}$ 为选择不同交通方式出行时所带来的显性效用; $m_{\text{mean } w}$ 表示选择 w 种方式出行的平均出行费用; $d_{\text{mean } w}$ 表示选择第 w 种方式出行的平均距离; \ln 表示居民收入水平; sd_1 、 sd_2 则表示哑变量优化赋值后的乘车满意度; $a_{0,1,2,3,4,5}$ 为显性效用方程系数。

从决策肢上下层结构可知, 每个具体方案的选择必然会优先触发上层选择概率, 而下层则是在上层因素被选择的情况下条件性的发生。根据相关模型的理论可得:

$$P = \frac{e^{\lambda_i (V_w + \varepsilon_i)}}{\sum_{i=1}^n e^{\lambda_i (V_w + \varepsilon_i)}}, (w = \text{bus, car, 非}) \quad (5)$$

4. 模型参数估计

极大似然求估法是一种基于观测变量分布明确或已知的前提下, 通过对似然函数或方程进行对数变换, 利用方程求导可积的特点来实现对原始方程降维, 进而得出相关参数最佳近似值的方法[11]。由于 NL 模型表达式为对数函数与多项式的结合体, 为了提高参数估计的准确性, 本文通过 SPSS 软件来获取质量较高的参数估计值。

5. SP 调查问卷

SP 调查全称为状态偏好调查(Stated Preference Survey), 该调查起初被用于了解消费者对于市场不同产品的喜好程度以及各种服务的满意度, 随着交通专家学者的深入研究, SP 调查渐渐地被引入到交通工程领域[12]。SP 的调查数据可以在历史趋势分析的基础上适当地根据调查者来制定参数水平范围[13], 此外 SP 调查对样方容量要求较少, 各个问题之间存在较好的关联性, 可以侧面验证和获取多种类型的数据[14], 因此 SP 具有较小的样本量且可以获取较多高效数据的特点, 更适合发起移动互联出行问卷调查活动[15]。

6. 实例分析

本章主要以 2019 年 1 月 SP 问卷调查为基础, 以桂林市七星区居民选择公交出行为例, 结合前文理

论分析与相关模型建立情况, 计算桂林市居民在出行前关于出行方式选择效用函数并求解算出选择行为所对应的概率函数。具体调查结果见下表 1 所示。

Table 1. A list of comprehensive information about households

表 1. 居民家庭综合信息一览表

| 家庭收入 | 人数 | 家庭人口总数 | 人数 | 家庭小汽车保有量 | 人数 |
|---------|-----|--------|-----|----------|-----|
| 0~5 万 | 301 | 1 人 | 7 | 0 辆 | 182 |
| 5~9 万 | 105 | 2 人 | 35 | 1 辆 | 294 |
| 9~13 万 | 84 | 3 人 | 224 | 2 辆 | 70 |
| 13~18 万 | 35 | 4 人 | 154 | 3 辆 | 21 |
| 18 万以上 | 56 | 5 人 | 77 | 4 辆 | 7 |
| | | 6 人 | 60 | 5 辆 | 7 |

根据 SP 调查数据, 低碳出行理念良好的人且愿意选择公交出门的人占比 38.25%, 此类居民的出行路线通常较为固定且对于出行效率要求不高; 对于出行效率要求和出行路线自由度较高的居民通常愿意选择小汽车出行, 其中占比 36.44%; 对于出行距离较短且经济条件较为普通的居民则多数以非机动车来作为自己的首选出行工具。相比其他交通工具, 非机动车受交通拥堵的影响较小, 且能够极易穿梭于巷子和街道, 为该类出行者提供抄近道的便利, 占比 25.31%。具体哑变量编码结果见下表 2 所示。

Table 2. List of dummy variable coding for qualitative indicators of travel mode

表 2. 出行方式定性指标哑变量编码一览表

| 原定性变量编码值 | 哑变量赋值 |
|----------------------|----------------------------|
| 舒适度 = 1 (车内宽松舒畅) | 所有 $SD = 0$ |
| 舒适度 = 2 (车内比较拥挤) | $SD_2 = 1$, 其他 $SD = 0$ |
| 舒适度 = 3 (车内非常拥挤) | $SD_3 = 1$, 其他 $SD = 0$ |
| 共享电动/单车资源 = 1 (资源充裕) | 所有 $FJD = 0$ |
| 共享电动/单车资源 = 2 (资源紧张) | $FJD_2 = 1$, 其他 $FJD = 0$ |
| 共享电动/单车资源 = 3 (资源缺乏) | $FJD_3 = 1$, 其他 $FJD = 0$ |

经 SPSS 软件求解可得, 效用函数的参数估计值与置信度结果总体上符合预计估计, 尽管有点局部参数置信水平收敛效果一般, 但任然处于可接受的范围内。具体 SPSS 回归分析参数标定结果见下表 3 所示。

Table 3. List of public transport utility function parameters calibration

表 3. 公交出行效用函数参数标定一览表

| 变量 | B 值 | Sig 值 |
|--------------|--------|-------|
| 平均出行距离(d/km) | 0.846 | 0.004 |
| 年收入情况(万元) | 1.252 | 0.050 |
| 出行费用(元) | -1.318 | 0.034 |
| 满意度 d_1 | 13.864 | 0.013 |
| 满意度 d_2 | 12.467 | 0.023 |

因此选择公交出行的效用选择函数可表示为:

$$V_{\text{bus}} = -12.255 - 1.318 * M_{\text{mean bus}} + 0.846 * d_{\text{mean bus}} + 1.252 * \text{Income} + 13.864 * SD_1 + 12.467 * SD_2 + 0 * SD_3 + \varepsilon_{\text{bus}} \quad (6)$$

综上所述, 桂林市居民选择公交出行的平均效用分为 9.6 分, 平均选择概率为 0.62, 模型求解结果大体符合预期估计, 具有一定的研究意义。

7. 结论

该研究细化优化各个哑变量分组分类的情况, 提高了模型的适用人群, 但本文只考虑方式并没有考虑到人对于时间和方式同时选择同时考虑的这一情况, 对此还需要开展进一步改善, 对于超长距离或超短距离出行方面考虑不够, 需要进一步研究和探讨。

本文在 Gumbel 假设分布的前提下, 通过对随机效用模型进行深入剖析, 在借鉴显性效用函数模型特点的基础上, 引入了哑变量并对模型性质指标进行分类再赋值, 在保证模型解释变量保质扩容的情况下, 建立了个体选择各交通方式出行的通项效用函数, 使得模型能够具有一定的普遍性。揭示了居民选择公交出行所衍生出的效用机制以及选择公交出行的不确定度。

参考文献

- [1] Jin, F.L., An, K. and Yao, E.J. (2020) Mode Choice Analysis in Urban Transport with Shared Battery Electric Vehicles: A Stated-Preference Case Study in Beijing, China. *Transportation Research Part A*, **133**, 95-108. <https://doi.org/10.1016/j.tra.2020.01.009>
- [2] Thorhauge, M., Swait, J. and Cherchi, E. (2020) The Habit-Driven Life: Accounting for Inertia in Departure Time Choices for Commuting Trips. *Transportation Research Part A*, **133**, 272-289. <https://doi.org/10.1016/j.tra.2020.01.023>
- [3] 米均. 考虑出行成本的城市居民出行方式选择研究[D]: [硕士学位论文]. 成都: 西南交通大学, 2015.
- [4] 孙海云. Logit 模型参数估计方法的研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2016.
- [5] Zhu, T.T., Long, J.C. and Liu, H.X. (2019) Optimal Official Work Start Times in Activity-Based Bottleneck Models with Staggered Work Hours. *Transportmetrica B: Transport Dynamics*, **7**, 657-683. <https://doi.org/10.1080/21680566.2018.1460881>
- [6] Yuan, Y.H., Raubal, M. and Liu, Y. (2012) Correlating Mobile Phone Usage and Travel Behavior—A Case Study of Harbin, China. *Computers, Environment and Urban Systems*, **36**, 118-130. <https://doi.org/10.1016/j.compenvurbsys.2011.07.003>
- [7] 刘玲玲, 刘巍巍, 董洁霜. 基于交叉巢式 logit 模型的交通出行方式选择研究[J]. 物流工程与管理, 2017, 39(6): 126-128, 125.
- [8] Masoumi, H.E. (2019) A Discrete Choice Analysis of Transport Mode Choice Causality and Perceived Barriers of Sustainable Mobility in the MENA Region. *Transport Policy*, **79**, 37-53. <https://doi.org/10.1016/j.tranpol.2019.04.005>
- [9] 李梦凡. 交通信息诱导下个体出行选择行为研究[D]: [硕士学位论文]. 合肥: 合肥工业大学, 2018.
- [10] 邵响泓, 王炜, 程琳. 出行方式决策的随机效用模型研究[J]. 公路交通科技, 2006(8): 110-115.
- [11] Mayo, F.L. and Taboada, E.B. (2019) Ranking Factors Affecting Public Transport Mode Choice of Commuters in an Urban City of a Developing Country Using Analytic Hierarchy Process: The Case of Metro Cebu, Philippines. *Transportation Research Interdisciplinary Perspectives*, **4**, 100078. <https://doi.org/10.1016/j.trip.2019.100078>
- [12] 唐立. 多元出行信息影响下的交通选择行为研究[D]: [博士学位论文]. 成都: 西南交通大学, 2015.
- [13] Bergantino, A.S., Capurso, M. And Hess, S. (2020) Modelling Regional Accessibility to Airports Using Discrete Choice Models: An Application to a System of Regional Airports. *Transportation Research Part A: Policy and Practice*, **132**, 855-871. <https://doi.org/10.1016/j.tra.2019.12.012>
- [14] 温惠英, 汤左溢. 基于 Nested Logit 与 Random Parameters Logit 模型的摩托车事故伤害估计与对比[J]. 华南理工大学学报(自然科学版), 2018, 46(11): 83-91.
- [15] 刘梦琪. 基于 Bootstrap 法的城市居民出行行为 Nested Logit 模型研究[D]: [硕士学位论文]. 成都: 西南交通大学, 2018.