



交通信息与 出行者路线选择

干宏程 著



科学出版社

014023287

U495
29

交通信息与出行者 路线选择

于宏程 著



科学出版社

北京



北航

C1710184

U495
29

内 容 简 介

对信息环境下的出行者决策行为进行建模是国际热点研究领域. 本书阐述离散选择模型的基本原理、常见模型、集计预测方法、常用检验方法以及实用建模技巧, 建立一系列先进的用于刻画交通信息下的出行者决策行为的离散选择模型, 给出模型应用案例. 所建模型有带序 Logit 面板模型、随机效应 Probit 面板模型、随机效应 Logit 面板模型、混合 Logit 模型、随机系数 Probit 模型等.

本书可供高等院校交通、经济、管理等相关专业的高年级本科生、研究生以及教师学习和参考, 也可供从事交通规划与管理的专业人士参考.

图书在版编目(CIP)数据

交通信息与出行者路线选择/干宏程著. —北京: 科学出版社, 2014.1

ISBN 978-7-03-039572-6

I. ①交… II. ①干… III. ①交通信息系统-决策模型 IV. ①U495

中国版本图书馆 CIP 数据核字 (2014) 第 008224 号

责任编辑: 徐园园 赵彦超 / 责任校对: 刘亚琦
责任印制: 赵德静 / 封面设计: 耕者设计工作室

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

骏杰印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2014 年 2 月第 一 版 开本: 720×1000 1/16

2014 年 2 月第一次印刷 印张: 10 3/4

字数: 200 000

定价: 48.00 元

(如有印装质量问题, 我社负责调换)

前 言

向出行者提供实时交通信息是智能化交通管理的重要手段. 信息的作用本质上取决于出行者对信息的响应 (如路径、交通方式、出发时刻的选择), 因此研究信息环境下的出行者决策行为非常重要. 研究成果有助于更高效合理地进行先进出行者信息系统的规划设计和运营管理.

信息环境下的出行者决策行为是一个极具挑战性的研究领域, 吸引了交通、经济、管理、行为科学、心理学等领域的科学家. 目前该领域主流的行为建模方法是基于随机效用理论的离散选择分析, 运用该方法建立的离散选择模型能刻画决策行为与出行者属性、出行特征、交通状况以及实时信息等因素的关系, 揭示信息对出行者决策的影响机理. 国际上该领域已取得不少研究成果, 但是日新月异的信息技术使该领域不断出现新难题. 我国学者在该领域的研究起步较晚, 理论研究水平与国际先进水平差距较大. 本人在出版此书前尚未见到国内有该领域的学术专著. 本书旨在为研究和实践人员从事该领域的理论和应用研究提供技术参考.

本书侧重介绍本人在信息环境下的出行者决策行为建模方面的主要研究成果, 其主要内容和特色如下:

(1) 从方法论的角度, 系统、完整、细致地阐述离散选择分析的理论背景和基本原理、常见离散选择模型以及集计预测方法, 建立一系列用于提高出行决策行为模型的性能的先进离散选择模型;

(2) 从模型应用的角度, 给出以我国大城市的先进出行者信息系统为载体的研究案例, 归纳建模过程中各种有效检验方法和实用技巧, 体现成果的良好应用价值.

作者的研究工作得到了国家自然科学基金 (No. 51008195)、上海市教育委员会科研创新基金 (No. 09YZ205)、上海理工大学引进人才启动

基金 (No. X693)、上海市重点学科建设项目 (No. S30504)、上海市一流学科建设计划 (No. S1201YLXK)、上海理工大学人文社科基地培育基金 (No. XZJ04)、国家科技攻关重大专项 (No. 2002BA404A08) 等的资助, 在此表示感谢.

本人在学术探索道路上得到了许多学者的帮助. 感谢我的硕士、博士生导师——教育部首届长江学者、同济大学孙立军教授将我引入科学研究的殿堂, 教我做人做学问的道理, 孙教授是我前行道路上的一盏明灯. 感谢同济大学陈建阳教授多年来对我的指导. 感谢美国的 Xin Ye 博士参与我的研究.

感谢科学出版社相关人员的辛勤付出. 感谢上海理工大学马良教授对本书出版的帮助. 感谢我的研究生苏红在本书编辑校对上的付出.

限于本人学识能力, 书中难免有纰漏和不当, 欢迎读者批评指正.

千宏程

2013 年 8 月

目 录

前言

第 1 章	绪论	1
第 2 章	个体选择行为理论	5
2.1	选择理论的框架	5
2.1.1	决策者	5
2.1.2	选择项	6
2.1.3	选择项的属性	6
2.1.4	决策规则	7
2.2	理性行为	7
2.3	经济消费理论	7
2.4	消费者理论的扩展	10
2.5	离散选择理论	11
2.6	概率选择理论	16
2.6.1	常效用方法	16
2.6.2	随机效用方法	21
第 3 章	二元选择模型	24
3.1	把随机效用理论变得有可操作性	24
3.1.1	二元选择情形	24
3.1.2	确定项和随机项	24
3.1.3	系统项的构造	25
3.1.4	干扰项的构造	27
3.2	常见二元选择模型	28
3.2.1	二元 Probit	28
3.2.2	二元 Logit	29

3.2.3	Probit 和 Logit 模型的极限情形	30
3.3	二元选择模型的案例	30
3.3.1	二元 Probit 模型	30
3.3.2	二元 Logit 模型	31
3.4	二元选择模型的最大似然估计	32
3.4.1	二元选择模型最大似然估计的一般性构造	32
3.4.2	最大似然估计的计算 (寻优算法)	34
3.4.3	最大似然估计在特定二元模型中的应用	36
3.5	最大似然估计结果示例	39
3.5.1	二元选择模型估计结果	39
3.5.2	一个二元选择模型的例子	41
第 4 章	多元选择模型	44
4.1	多元选择理论	44
4.2	多元 Logit 模型	46
4.2.1	多元 Logit 定义	46
4.2.2	Gumbel 分布: 基本特征	47
4.2.3	多元 Logit 推导	48
4.2.4	多元 Logit 模型的极端情形	49
4.2.5	线性参数 Logit 模型	50
4.3	Logit 模型的特征	50
4.3.1	IIA 特性	50
4.3.2	Logit 的弹性	52
4.3.3	增量 MNL 模型	54
4.4	多元 Logit 模型的构造	55
4.5	多元 Logit 模型的估计	56
4.5.1	最大似然估计	56
4.5.2	针对重复观测的最大似然估计	58
4.6	估计结果的举例	59

4.7 其他多元选择模型	61
4.7.1 随机系数 Logit	61
4.7.2 带序 Logistic 模型	62
4.7.3 广义极值模型	62
4.7.4 多元 Probit 模型	65
第 5 章 二元 Logit 模型	66
5.1 建模任务	66
5.2 建模方法	66
5.3 模型应用实例	67
5.3.1 数据	67
5.3.2 模型估计结果以及讨论	68
第 6 章 二元 Probit 模型	71
6.1 建模任务	71
6.2 建模方法	71
6.3 模型应用实例	72
6.3.1 数据	72
6.3.2 模型估计结果以及讨论	72
第 7 章 多元 Logit 模型	76
7.1 建模任务	76
7.2 建模方法	76
7.3 模型应用实例	77
7.3.1 数据	77
7.3.2 模型估计结果及讨论	78
第 8 章 二元 Logit 面板模型	81
8.1 建模任务	81
8.2 建模方法	81
8.3 模型应用实例	83
8.3.1 数据	83

8.3.2	模型估计结果以及讨论	83
第 9 章	多种面板数据模型比较	86
9.1	建模任务	86
9.2	建模方法	86
9.2.1	基于随机效应的模型	86
9.2.2	基于随机系数的模型	88
9.2.3	基于随机效应和随机系数的模型	89
9.3	模型应用实例	90
9.3.1	数据	90
9.3.2	模型估计结果及讨论	90
第 10 章	二元 Probit 面板模型	94
10.1	建模任务	94
10.2	建模方法	94
10.2.1	横截面模型	94
10.2.2	面板模型	95
10.3	模型应用实例	96
10.3.1	数据	96
10.3.2	模型估计结果及讨论	97
第 11 章	带序多元 Probit 面板模型	100
11.1	建模任务	100
11.2	建模方法	100
11.2.1	横截面带序 Probit 模型	100
11.2.2	面板数据带序 Probit 模型	101
11.3	模型应用实例	103
11.3.1	数据	103
11.3.2	模型估计结果及讨论	104
第 12 章	二元随机系数 Probit 模型	108
12.1	建模任务	108

12.2	建模方法	108
12.3	模型应用实例	110
12.3.1	数据	110
12.3.2	模型估计结果及讨论	111
第 13 章	集计预测技术	115
13.1	集计问题	115
13.2	集计方法分类	116
13.3	集计方法的描述	117
13.3.1	平均化个体方法	117
13.3.2	分类法	119
13.3.3	统计微分法	121
13.3.4	直接积分法	123
13.3.5	样本枚举法	125
13.4	各种集计预测方法的比较	127
第 14 章	建模中的常用检验及实用技巧	131
14.1	引言	131
14.2	建模的艺术	131
14.3	一个方式选择模型案例	132
14.4	变量构造的检验	133
14.4.1	系数估计值的非正式检验	133
14.4.2	渐近 t 检验的使用	134
14.4.3	同时含多个参数的置信区间	135
14.4.4	似然率检验的使用	136
14.4.5	拟合优度指标的使用	137
14.4.6	共通属性的检验	138
14.4.7	非嵌套假设的检验	138
14.4.8	非线性构造的检验	139
14.4.9	受限的估计	140

14.5	模型结构的检验	141
14.5.1	IIA 假设的检验	142
14.5.2	口味差异性的检验	144
14.5.3	异方差性的检验	146
14.6	预测检验	147
14.6.1	离群值分析	148
14.6.2	市场分块预测检验	148
14.6.3	政策预测检验	150
参考文献		152
索引		158

第1章 绪 论

智能交通系统 (ITS) 自从 20 世纪 80 年代在发达国家被提出, 经过 20 余年的研发, 已在全球范围内得到了普及, 提高了交通管理效率, 改善了出行者的出行质量. 先进出行者信息系统 (ATIS) 是 ITS 的重要组成部分, 其建设在我国正如火如荼地开展. ATIS 通过可变信息标志 (VMS)、车载导航装置、交通广播、智能手机等向出行者提供实时的出行相关信息, 帮助他们快速、便捷、经济地完成出行. ATIS 的效果本质上取决于出行者对信息的响应行为 (如路径、交通方式、出发时刻的选择), 因此研究信息环境下的出行者决策行为至关重要. 研究成果有助于正确理解出行者在提供信息条件下的出行行为, 科学高效地进行 ATIS 的规划设计和运营管理.

国际上, 信息环境下的出行者决策行为自 ITS 概念被提出以来一直是个热点研究领域^[1-42]. 由于人类行为高度复杂, 该领域的探索极具挑战性, 吸引了交通科学、经济学、行为科学、管理科学、心理学等领域的科学家. 该领域的研究主题非常丰富. 国外曾有学者归纳了该领域中一系列相互关联的代表性问题: ①如果有实时的信息可使用, 那么哪些类型的出行者会使用信息? 人们使用信息用于什么类型的出行? ②出行者对信息会做出怎样的反应? 信息如何直接影响正在或即将出行的出行者的决策? 信息如何影响出行的情境? ③出行者喜欢哪些类型的信息? ④出行者愿意为信息支付多少费用? ⑤出行者如何评价他们获取信息并对信息做出响应后的获益? ⑥出行者对信息带给他们的好处的评价如何影响今后他们对以上问题的回答?

尽管该领域的文献报道已经不少, 但是人们对出行者的信息响应行为的认识至今还很局限. 例如, 谁都不能准确预测, 特定路网中某个特定位置处的一块图状路径信息板^[43,44](graphical route information panel,

GRIP)(图 1.1) 发布的某条特定信息会对驾驶员个体以及整个路网造成多大的影响. 以往研究大多针对上下班通勤者等特定类型的出行者, 对公交车驾驶员、商用车司机等非通勤者的研究较少. 学者们提出的出行决策行为模型往往理论性较强, 计算复杂, 而实用性欠佳. 能够评价路网层面的 ATIS 效果的模型还很不完善, 尚缺少一般性模型. 学术界任重而道远.



图 1.1 上海的 GRIP

信息环境下出行者决策行为研究的两个关键环节是行为数据获取和行为分析与建模. 本书主要侧重行为分析与建模.

出行行为分析方法主要有描述性统计、数理统计方法和离散选择分析. 描述性统计主要是对行为数据进行简单处理, 用以推断出行者出行决策的标准和这些标准的相对重要性. 常见的数理统计方法有主成分分析法、相关分析法、多元回归、聚类分析等.

离散选择分析属于计量经济学的范畴, 基于随机效用理论, 具有较强的数学性, 适用于刻画个体决策行为^[45-48]. 它是目前对信息环境下的出行者决策行为进行建模的主流方法. 运用该方法得到的离散选择模型 (discrete choice model) 可以刻画出出行者决策与出行者个体特征、出行特征、交通状况以及信息等因素之间的关系, 揭示信息对出行者决策的影响机理. 研究中最常用的模型是多元 Logit 模型 (multinomial Logit model, MNL). 然而, MNL 的最大缺陷是其本身的不相关选择项之间相互独立的特性 (IIA 特性), 为此学者们提出分层 Logit 模型 (nested Logit

model) 和多元 Probit 模型等其他模型来克服 MNL 的不足. 此外, 学者们为了考虑出行者偏好差异 (随机口味) 及选择行为在时间序列上的相关性 (如对同一受试者做多次观测的情形, 对日复一日学习过程进行建模的情形), 又提出混合多元 Logit 模型 (mixed multinomial Logit model, MML). 还有其他离散选择模型在研究中被不断地提出和应用^[49].

运用离散选择分析方法对信息环境下的出行者决策行为进行建模, 是国际上的一个热点领域, 相关研究成果二十多年来在交通科学领域的国际权威期刊上一直有报道. 我国学术界对这一热点领域的探索工作起步较晚, 总体研究水平与国际先进水平之间有较大差距. 虽然随着 ITS 建设加快, 我国学者在这一领域的研究越来越活跃, 但当前的研究从方法论上来说大多以应用研究为主. 此外, 目前尚未见到国内有信息环境下的出行者决策行为离散选择建模的学术专著. 国内 ITS 的大力建设正热切期盼着学术界能出现这方面的学术专著, 可用于指导 ATIS 实践. 本书的出版顺应了上述需求.

Ben-Akiva 等于 1985 年发表的著作 [45] 是交通领域一本经典的关于离散选择模型的书. 本书讲解离散选择分析的基础知识时借鉴了该书.

本书报道了作者在信息环境下的出行者决策行为建模方面的主要研究成果, 其主要内容和特色如下.

(1) 系统、完整、细致地阐述离散选择分析的理论背景和基本原理、常见离散选择模型、集计预测方法以及建模中的常用有效检验方法和实用技巧. 这些知识对离散选择模型的初学者而言是非常有用的入门知识, 安排在第 2~4 章、第 13 章、第 14 章.

(2) 针对离散选择模型家族中的二元 Logit、多元 Logit、二元 Probit 等常见模型, 给出以我国大城市的先进出行者信息系统为载体的研究案例, 使读者对常见模型的应用有初步的实战经验. 这些内容安排在第 5~7 章.

(3) 针对信息环境下的出行者决策行为的特点, 本书建立一系列用于提高出行决策行为模型的性能的先进离散选择模型. 这些模型包括多种

二元 Logit 面板模型、二元 Probit 面板模型、带序多元 Probit 面板模型、二元随机系数 Probit 模型. 这些建模方法对离散选择模型的初学者而言是具有一定难度的高级知识. 对于每种建模方法, 本书给出以我国大城市的先进出行者信息系统为载体的研究案例, 体现建模方法的良好应用价值. 这些内容安排在第 8~12 章.

第2章 个体选择行为理论

本章首先介绍选择理论的一般性框架以及共同假设, 然后介绍经典的经济消费者理论和离散选择理论, 接着介绍概率选择模型的理论及特点. 概率选择模型是建立经验性离散选择模型的基础.

2.1 选择理论的框架

选择可视为一个连续的决策过程的一个结果. 该决策过程包括以下步骤: 定义选择问题 → 生成选择项 → 评价选择项的属性 → 选择 → 实施所做的选择.

一种特定的选择理论是一种方法的集合, 该集合定义了如下的要素: 决策者; 选择项; 选择项的属性; 决策规则.

要注意的是, 并非所有观测到的选择行为都是通过以上过程做出的. 例如, 某些人会按习惯、照直觉行事, 或者模仿某一个他认为的专家或领导者的行为. 但是, 这种决策行为可以表示为决策者只产生一个选择项的选择过程.

2.1.1 决策者

决策者可以是一个个体、一组人 (如家庭)、一个组织 (如公司). 不同个体所面临的选择情形不同, 口味也不同. 因此, 尽管我们最终关心的是预测集计需求, 但是必须明确地对待不同个体在选择过程中的差异. 例如, 交通方式选择案例中的小汽车出行成本, 取决于车辆类型和当地油价. 此外, 愿意支付更高的小汽车出行费用的程度可能依赖于个体的收入.

此外, 口味差异可能在群体决策过程中体现, 因为群体内部的交互作

用的差异会影响结果. 例如, 对于买车的决策, 有些家庭是一个主导地位的人说了算, 而有些家庭则需要多个成员热烈讨论商量后才能定.

2.1.2 选择项

根据决策者所处的环境, 我们可以确定一个称作“总选项集”的集合. 单个决策者仅考虑总选项集的一个子集, 称为一个“选项集”. 选项集中的选项对于决策者而言是可行的, 并且在决策过程中是已知的. 可行性是由诸多约束条件决定的. 这些约束条件包括: 物理上的可利用性 (如在通勤出行者的家庭与工作地点之间是否有公交车路线)、可利用的资金资源 (如低收入工作者很难支付得起出租车费用)、时间上的可行性 (如长距离通勤出行者采用步行方式是不现实的)、信息掌握程度的局限性 (如缺乏有关公交服务的信息).

要区分两类选项集: 连续的选择项、非连续的选择项.

连续选项集在新古典主义微观经济学的需求分析中很常见. 例如, 一个家庭将一笔可支配收入 (I) 用在购买牛奶 (购买量为 q_1 , 价格为 p_1)、面包 (购买量为 q_2 , 价格为 p_2) 以及黄油 (购买量为 q_3 , 价格为 p_3) 三种商品上, 相应的选项集可以由 $p_1q_1 + p_2q_2 + p_3q_3 = I$ 这一关系式表示.

非连续选项集的一个例子为: 某个家庭要购买三款电视机中的某一款. 三款电视机为 A, B, C . 因此, 选项集仅仅是一个点集 $\{A, B, C\}$.

2.1.3 选择项的属性

选择项吸引力程度可用一个属性值向量来评价. 属性值的度量可以是序次的 (如小汽车是最快捷的模式), 也可以是数值的 (如出行费用为 50 元). 如果选择项是均质的 (如牛奶或黄油), 并且一个选择项就是这些货物数量的一个向量, 那么属性向量仅仅就是数量. 但是, 本书关心的是这样的情景: 选择项是不均质的, 不同决策者可能面临不同的选择集, 评价不同的属性, 对同一选择项的同一属性的看重程度可能不同. 对于这些情景, 直接把每个选项按照其属性进行一般性的特征化 (而非处理选项的数量) 就更为自然了.