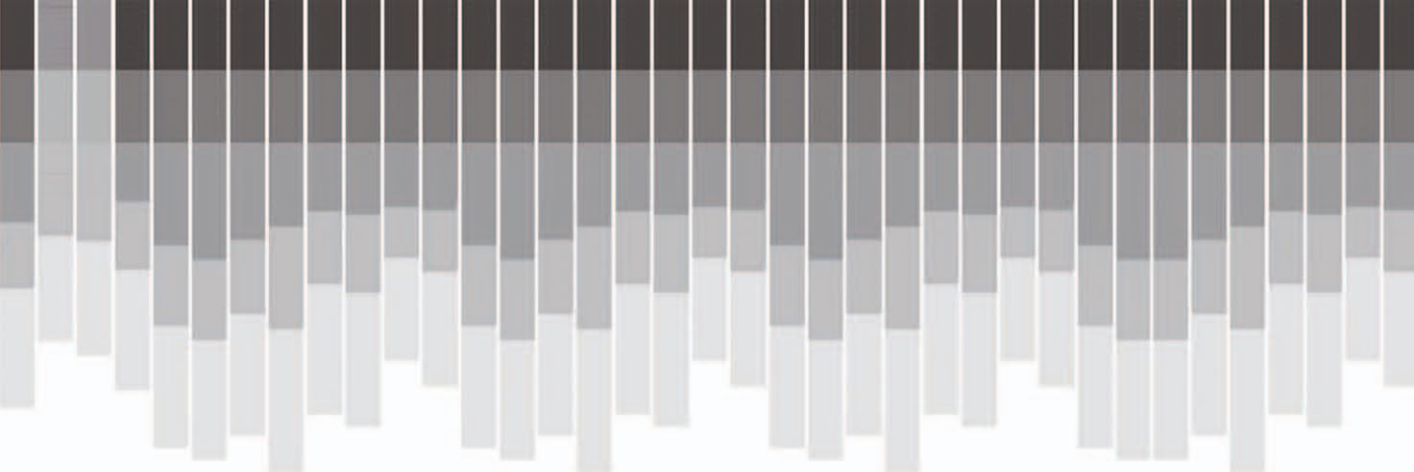




出行行为分析的 高级计量经济学方法和应用

干宏程 著



同济大学出版社
TONGJI UNIVERSITY PRESS

出行行为分析的 高级计量经济学方法和应用

干宏程 著



同济大学出版社
TONGJI UNIVERSITY PRESS

内 容 提 要

出行者行为的计量经济学模型研究是针对集交通、经济、管理、行为科学、心理学等学科为一体的交叉领域。本书秉持“理论需联系实践、知识应融会贯通”之观念,完整、细致地阐述混合 Logit 模型、广义线性混合模型和广义估计方程这三种交通领域先进的出行行为计量分析方法的基本原理和技术细节,并给出多个针对当前学术热点的研究案例,例如,“互联网+”环境下的交通方式选择行为、可变信息标志影响下的路线选择行为、电动汽车购买意愿、停车换乘行为、行程时间可靠性对路线选择的影响等。

本书可供高等院校交通、经济、管理、行为科学等相关专业的研究生、高年级本科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

出行行为分析的高级计量经济学方法和应用/干宏程著. —上海:同济大学出版社,2017. 10
ISBN 978-7-5608-7414-2

I. ①出… II. ①干… III. ①计量经济学—应用—交通分析 IV. ①U491. 1

中国版本图书馆 CIP 数据核字(2017)第 222936 号

出行行为分析的高级计量经济学方法和应用

干宏程 著

责任编辑 陆克丽霞

责任校对 徐春莲

封面设计 陈益平

出版发行 同济大学出版社 www.tongjipress.com.cn
(地址:上海市四平路 1239 号 邮编:200092 电话:021-65985622)
经 销 全国各地新华书店
排 版 南京月叶图文制作有限公司
印 刷 常熟市大宏印刷有限公司
开 本 787 mm×1 092 mm 1/16
印 张 6.25
字 数 156 000
版 次 2017 年 10 月第 1 版 2017 年 10 月第 1 次印刷
书 号 ISBN 978-7-5608-7414-2

定 价 36.00 元

本书若有印装质量问题,请向本社发行部调换 版权所有 侵权必究

前 言

出行者行为的计量经济学模型研究是针对集交通、经济、管理、行为科学、心理学等学科为一体的交叉领域,它也是非集计出行需求预测的理论基础。来自计量经济学的离散选择模型是该领域的主流研究方法。目前,离散选择模型家族中的混合 Logit 模型代表了交通领域出行行为分析的先进水平。此外,近年来,统计学中的广义线性混合模型和广义估计方程等先进方法也被应用于出行行为研究中,并取得一定的成功。笔者在撰写此书前尚未见到国内交通领域有著作或教材系统地讲解混合 Logit 模型、广义线性混合模型和广义估计方程等先进的出行行为分析方法。因此,本书旨在填补这一空白,为从事该领域的理论和应用研究人员提供理论指导和技术参考。

笔者从事出行行为探索已有十多年,取得了一些具有应用价值的研究成果。由于时间、精力及篇幅所限,本书仅收纳了其中的一部分最新成果。本书的主要内容和特色如下:

(1) 从方法论的角度,完整、细致地阐述混合 Logit 模型、广义线性混合模型以及广义估计方程的基本原理和技术细节。

(2) 从方法应用的角度,给出多个针对当前学术热点的研究案例,以及在计量分析软件中的实现过程,增强了读者的“实战”感受。

本书的研究工作得到了国家自然科学基金(No. 51008195)、上海市曙光学者人才资助项目(No. 15SG41)、上海市自然科学基金项目(No. 15ZR1428100)、同济大学道路与交通工程教育部重点实验室开放基金(No. K201302)、上海市重点学科/一流学科/高原学科“管理科学与工程”建设项目



(No. S30504、No. S1201YLXK)、上海教委科研创新基金(No. 09YZ205)、上海理工大学引进人才启动基金(No. X693)、沪江社科基金重点项目(No. 14XSZ02)、上海理工大学人文社科攀登计划资助项目(No. 16HJPD—A06)、国家科技攻关重大专项(No. 2002BA404A08)等项目的支持。

本书的部分研究工作是笔者在美国西北大学工作学习期间完成的,感谢在此期间西北大学的导师、交通/管理学科双聘教授 Hani Mahmassanni 提供的工作环境和有益指导。美国 Xin Ye 博士、香港大学柏杨博士以及笔者所带的研究生苏红、赵亚菲、黄冰等参与了书中的部分研究工作,在此一并表示感谢。感谢所有对此研究给予过笔者指导和帮助的人。

感谢同济大学出版社陆克丽霞编辑在本书出版过程中的辛勤付出。感谢我的研究生孙亦凡在本书成稿过程中在编辑校对上的帮助。

感谢我的父母、妻子、岳父母等家人对我的研究工作的理解、支持和默默付出。

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。

2017 年 5 月

目 录

前言

第 1 章 绪论	1
第 2 章 混合 Logit 模型	4
2.1 选择概率	4
2.2 随机系数	6
2.3 误差-组成部分	7
2.4 取代型式	8
2.5 对任何随机效用模型的近似	9
2.6 仿真	10
2.7 面板数据	11
第 3 章 研究实例 1——多交通方式实时信息对出行行为的影响	13
3.1 背景	13
3.1.1 研究对象	13
3.1.2 数据	14
3.2 模型构造	14
3.3 模型讨论	15
3.4 模型应用	17
第 4 章 研究实例 2——可变信息标志行程时间信息对出行行为的影响	19
4.1 背景	19
4.1.1 研究对象	19
4.1.2 数据	19
4.2 模型构造	21
4.2.1 基于随机效应的模型	21
4.2.2 基于随机系数的模型	22



4.2.3 基于随机效应和随机系数的模型	23
4.3 模型讨论	24
4.3.1 模型估计结果	24
4.3.2 模型系数讨论	24
4.3.3 不同模型的系数讨论	25
第5章 研究实例3——电动汽车购买意愿研究	27
5.1 背景	27
5.1.1 研究对象	27
5.1.2 数据	27
5.2 数据分析和建模	28
5.2.1 模型构造	28
5.2.2 SAS软件分析过程	28
5.3 模型估计结果讨论	33
5.3.1 模型估计结果	33
5.3.2 模型讨论	35
5.4 模型应用	37
第6章 研究实例4——停车换乘行为研究	40
6.1 背景	40
6.1.1 研究对象	40
6.1.2 数据	40
6.2 数据分析和建模	42
6.2.1 样本基本分布	42
6.2.2 模型构造	42
6.2.3 SAS软件分析过程	42
6.3 模型估计结果讨论	46
6.3.1 模型估计结果	46
6.3.2 模型讨论	48
第7章 基于广义线性混合模型的出行行为分析	51
7.1 背景	51
7.2 数据	51
7.3 广义线性混合模型的建立	52



7.3.1 模型原理	52
7.3.2 候选模型	53
7.4 模型结果及讨论	55
7.4.1 模型估计结果	55
7.4.2 讨论	56
7.5 小结	57
第8章 广义估计方程在交通行为分析中的应用	58
8.1 广义估计方程基本原理	58
8.1.1 简介	58
8.1.2 理论基础	58
8.2 广义估计方程在 SPSS 中的实现	62
8.3 应用案例一	72
8.3.1 案例背景	72
8.3.2 模型估计结果	73
8.3.3 模型应用	75
8.4 应用案例二	78
8.4.1 案例背景	78
8.4.2 模型估计结果	79
参考文献	83

第1章 绪 论

在过去的半个多世纪里,以离散选择分析为基础的非集计出行需求模型研究应用在欧、美、日等发达国家和地区得到蓬勃的发展。来自计量经济学领域的离散选择模型通过对微观层面(个人、家庭、公司等)出行行为的分析,使得出行需求预测的质量、精度和应用价值较传统的集计出行需求模型得到了显著的提高。我国的交通领域至少从20世纪90年代开始应用离散选择模型进行出行需求预测的研究与实践。目前,简单易用的标准Logit模型是研究与实践人员使用最多的离散选择模型。然而,标准Logit模型的IIA特性也使其具有明显的缺陷。例如,不能考虑“随机口味”“跨期间重复观测结果的相关性”“受限制的取代型式”。为此,早在20世纪70年代,国外学者便开始致力于能克服标准Logit模型的部分或全部缺陷的其他离散选择模型的研究,例如,HEV模型、多元Probit模型、混合Logit模型。如今,在计算机速度和仿真手段快速进步的背景下,混合Logit模型以其可以近似几乎任何一种随机效用模型的优势成了广受学术界青睐的离散选择模型,代表了离散选择模型当前的先进水平。此外,也有学者将统计学领域的广义线性模型(GLM)、广义线性混合模型(GLMM)、广义估计方程(GEE)等先进的建模方法引入交通需求预测中,取得了明显优于标准Logit模型的效果。

然而,据笔者了解,国内交通领域虽然已有一些学者先后发表了运用混合Logit模型、GLMM、GEE等先进方法进行出行行为建模的文章,但是迄今尚未有已出版的、系统讲解这些先进的出行行为建模方法的专著或教材。这一现象实际上从一个侧面反映了我国在出行需求的计量经济学模型这一研究领域与国际先进水平的差距,这也着实令人遗憾。此外,笔者结合多年的教学经验发现,对混合Logit等建模方法的学习一知半解(甚至错误理解)和囫圇吞枣,或因畏难和求“速成”心理作祟而浅尝辄止的现象在交通领域的研究生中普遍存在。一本系统讲解混合Logit等先进的出行行为计量经济学方法的书或许可以成为研究人员较快地掌握出行行为分析的“先进武器”。

随着我国经济、文化水平的提高,以及“互联网+交通”的到来,交通需求多元化、变化快的特点愈发明显,整个社会对交通需求预测能力提出了更高的需求。交通基础设施规划、交通政策制定、综合运输体系规划等对先进的出行需求建模方法的依赖越来越强。

在以上背景下,本书秉持“理论需要联系实践”“知识应融会贯通”的观念,全面细致地阐述混合Logit模型、广义线性混合模型、广义估计方程这三种先进的出行行为分析方法,并给出多个应用这些方法开展学术热点研究的案例。



本书的内容组织如下:第1章为绪论。第2章系统地讲解混合 Logit 的相关理论知识,为初学者打下理论基础,第3章至第6章给出了4个混合 Logit 模型的应用案例,分别是“多交通方式实时信息对出行决策的影响”“可变信息标志行程时间信息对路线选择的影响”“电动汽车购买意愿”“停车换乘行为”。其中,第2章至第6章在本书中归为基础篇。第7章讲解广义线性混合模型,并给出它在“行程时间波动性对路线选择影响”研究中的应用案例。第8章讲解广义估计方程,并给出它在“多交通方式实时信息对出行决策的影响”和“可变信息标志对路线选择的影响”两项研究中的应用案例。第7章和第8章在本书中归为拓展篇。

图 1.1 给出了本书的内容与构架。

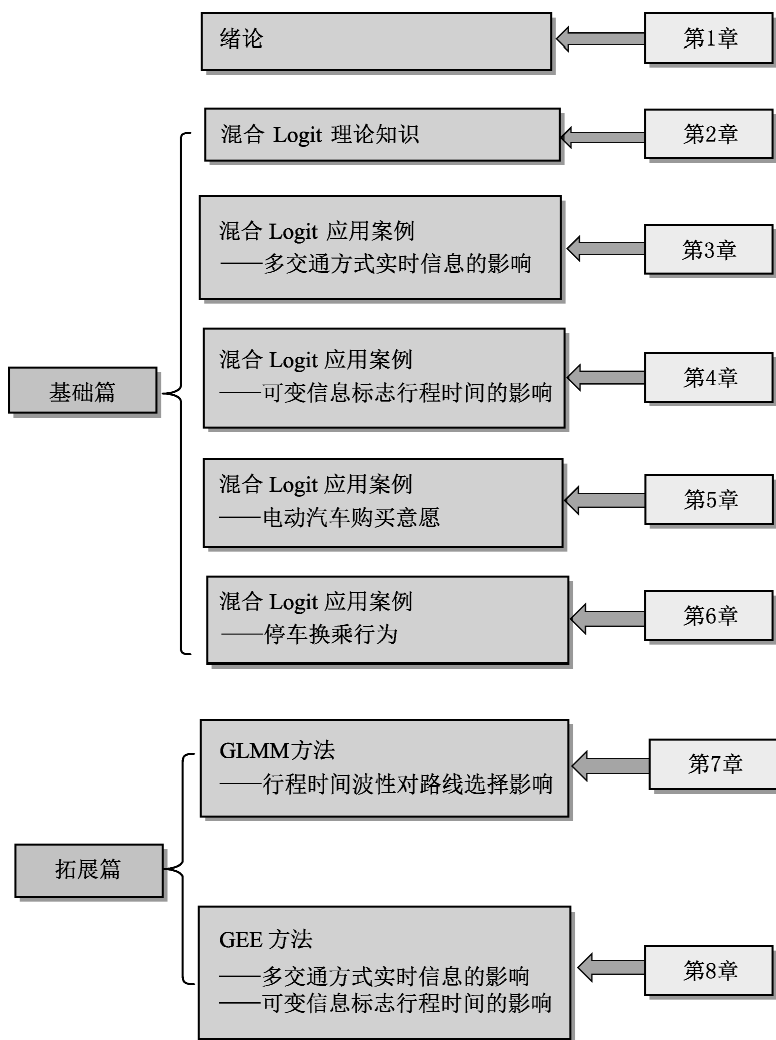


图 1.1 本书的内容与构架



本书具有两个特点:每一章实际上都是相对独立的,读者可以结合兴趣和知识背景选读部分章节;在案例研究的章节中,结合具体的计量分析软件给出了建模方法在软件中的具体实现过程,大大增强了读者身临其境的“实战感受”,极大地方便了读者的学习与实践。

第2章 混合 Logit 模型

2.1 选择概率

混合 Logit 模型是一种可以近似几乎任何一种随机效用模型的非常灵活的模型 (McFadden, Train, 2000)。它通过允许存在“随机口味差异”“不受限制的取代型式”“不可观察因素在时间上的相关性”这三个方面特征,从而解决了标准 Logit 模型的三个缺陷。不像 Probit 模型,它不受正态分布假设的约束,推导简单,选择概率的仿真在计算机上很简单。

混合 Logit 模型像 Probit 模型一样已被人们熟知很多年了,但是直至仿真技术的出现才被真正应用起来。Boyd, Mellman(1980)和 Cardell, Dunbar(1980)最早将它应用于小汽车需求建模中。上述的 Boyd 和 Cardell 等学者的研究中,解释变量不随决策者而改变,观测到的因变量是市场份额,而不是消费者个体的选择。因此,混合 Logit 模型涉及的烦琐积分在上述的 Boyd 和 Cardell 等学者的研究中,只需要针对整个市场做一次即可。Train 等(1987)和 Ben-Akiva 等(1993)把混合 Logit 模型应用到消费者个体层面的行为数据分析中,但也仅仅考虑了一维或二维的积分,求解起来较容易。计算机运行速度的提高以及人们对仿真方法认识的加深,使得混合 Logit 模型的威力得以充分展现。混合 Logit 模型的威力在 Bhat(1998a)、Brownstone, Train(1990)关于横截面数据的研究,以及 Erdem(1996)、Revelt, Train(1998)、Bhat(2000)关于面板数据的研究中得以证实。

混合 Logit 模型可以在许多不同的行为构造方式下推导出来,不同的推导提供了不同的解释。混合 Logit 模型是根据选择概率的函数形式来定义的。任何一种行为构造,只要其推导出的选择概率采用了这种特定形式的函数形式,都被称为混合 Logit 模型。

混合 Logit 模型的选择概率可以表达为如下形式:

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta \quad (2.1)$$

式中 $L_{ni}(\beta)$ ——在参数 β 上计算得到的 Logit 概率,即

$$L_{ni} = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^J e^{V_{nj}(\beta)}} \quad (2.2)$$



$f(\beta)$ ——密度函数。 $V_{ni}(\beta)$ 是效用的一部分,依赖于参数 β ,如果效用关于参数 β 是线性的,则 $V_{ni}(\beta) = \beta' x_{ni}$ 。这一条件下,混合 Logit 模型的选择概率为:

$$P_{ni} = \int \left[\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \right] f(\beta) d\beta \quad (2.3)$$

混合 Logit 模型的概率是不同 β 值处的 Logit 模型概率的加权平均值,其中权重为密度函数值 $f(\beta)$ 。在统计学文献中,若干函数的加权平均被称作一个混合函数,用做权重的密度函数被称作混合分布。混合 Logit 模型就是 Logit 函数在不同 β 取值处的混合,而混合分布为 $f(\beta)$ 。

标准 Logit 模型是混合 Logit 模型的一种特例,即在 $\beta = b$ 时 $f(\beta)$ 为 1,在 $\beta \neq b$ 时 $f(\beta)$ 为 0。式(2.3) 于是变为:

$$P_{ni} = \frac{e^{b' x_{ni}}}{\sum_j e^{b' x_{nj}}} \quad (2.4)$$

混合分布 $f(\beta)$ 可以是离散的,即 β 取有限个数的不同值。假设 β 可取 M 个值,分别记为 $b_1, b_2, \dots, b_M$,相应的取值概率为 $s_m, m = 1, 2, \dots, M$ 。这种情形下,混合 Logit 模型变成了心理学和市场营销学科中非常流行的“潜类别模型”。潜类别模型的应用实例有 Kamakura, Russell(1989)以及 Chintagunta, Jain, Vilcassim(1991)开展的工作。此种情况下的选择概率为:

$$P_{ni} = \sum_{m=1}^M s_m \left[\frac{e^{b'_m x_{ni}}}{\sum_j e^{b'_m x_{nj}}} \right] \quad (2.5)$$

以上构造对于诸如整个人群可以分为 M 个子群,每个子群有特定的选择行为或偏好这种情况非常有用。每个子群 m 在总人群中占的份额是 s_m 。 s_m 可以被研究人员连同所有的 b_m 估计出来。

在大多数所谓的混合 Logit 模型应用研究中, $f(\beta)$ 都被构造为连续函数。例如,假设 $f(\beta)$ 服从正态分布时,若均值为 b ,协方差为 ω ,则选择概率变为:

$$P_{ni} = \int \left[\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \right] \phi(\beta | b, \omega) d\beta \quad (2.6)$$

其中, $\phi(\beta | b, \omega)$ 是服从均值为 b 和协方差为 ω 的正态密度函数。研究人员要估计的是 b 和 ω 。对数正态分布、正态分布、三角形分布、伽马分布或任何其他分布都可以被使用。通过恰当地定义解释变量或密度函数,研究人员可以用混合 Logit 模型来表示任何一种效用最大化行为以及许多形式的非效用最大化行为。

有学者[如 McFadden, Train(2000)和 Chesher, Santos-Silva(2002)]提出了是否需要使用非退化混合分布(non-degenerative mixing distribution)以及某种分布形式是否



恰当的检验方法。

有个术语问题需要澄清。混合 Logit 模型中有两类参数。一类是 β , 它们的密度函数为 $f(\beta)$ 。第二类是描述密度函数的参数。以正态分布为例, 如果 β 服从均值为 b 、协方差为 w 的正态分布, 则 b 和 w 是 $f(\beta)$ 的参数。通常(并非总是), 研究人员关心的是估计 $f(\beta)$ 的参数。

把描述 $f(\beta)$ 的参数记为 θ 。那么 $f(\beta)$ 的更适合的记法应该是 $f(\beta | \theta)$ 。混合 Logit 模型选择概率 $P_{ni} = \int L_{ni}(\beta) f(\beta | \theta) d\beta$ 是 θ 的函数, 而不依赖 β 的值。参数被积分掉了。在这种意义下, β 就像 ϵ_{nj} 一样, 都被作为随机项被积分掉了。

在某些推导形式下, 混合 Logit 模型中的 β 取值可以表示决策者个人的口味。于是, 研究者可能想要获得每一位抽样的决策者的 β 信息以及描述 β 分布情况的 θ 。接下来, 本书将会描述如何使用经典的估计方法估计和解释 θ 。通过学习相关文献, 读者还可以掌握如何根据 θ 的估计结果以及每个决策者的选择结果来得到每个决策者的 β 信息的方法, 掌握同时得到 θ 信息和每个决策者的 β 信息的贝叶斯方法。

2.2 随机系数

可以用许多方法从效用最大化行为来推导混合 Logit 模型概率, 这些方法实际上是相同的, 只是给出的解释不同而已。最简单直观的, 同时也是在近期应用研究中广泛使用的推导方法是基于随机系数的推导方法。决策者面临 J 个选项。个体 n 从选项 j 中获得的效用构造为:

$$U_{nj} = \beta'_n x_{nj} + \epsilon_{nj} \quad (2.7)$$

式中 x_{nj} ——和选项及决策者有关的可观测变量;

β'_n ——与可观测变量对应的代表个体口味的系数向量;

ϵ_{nj} ——服从相同独立的极值分布的随机项。

系数在整个人群中依决策者不同而变化, 其密度函数为 $f(\beta)$ 。 $f(\beta)$ 是诸如 β 的均值和协方差等参数的函数。这种构造除了 β 随着决策者不同而变化这一特征外, 其余与标准 Logit 模型是相同的。

决策者知道他自己的 β_n 值和所有选项的 ϵ_{nj} , 并且在 $U_{ni} > U_{nj} \forall j \neq i$ 时选择选项 i 。研究者只能观测 x_{nj} , 但无法观测 β_n 或 ϵ_{nj} 。如果研究者观测到了 β_n , 那么选择概率就是标准 Logit 模型(因为 ϵ_{nj} 是 iid 极值分布)。也就是说, 依赖于 β_n 的条件概率为

$$L_{ni}(\beta_n) = \frac{e^{\beta'_n x_{ni}}}{\sum_j e^{\beta'_n x_{nj}}} \quad (2.8)$$

但是, 研究者并不知道 β_n , 因此不能以 β 为条件求条件概率。于是, 非条件选择概率应该是 $L_{ni}(\beta_n)$ 在所有可能的 β_n 取值上的积分, 公式见式(2.3)。



研究人员构造了许多用于描述系数的分布,其中用得最多的是正态分布和对数正态分布,即 $\beta \sim N(b, w)$ 和 $\ln(\beta) \sim N(b, w)$, b 和 w 是需要估计的参数。对数正态分布对于系数在所有决策中都取相同符号的情况非常有用。也有学者使用三角形分布和均匀分布。对于均匀分布密度函数, β 均匀地在 $b-s$ 和 $b+s$ 之间分布,其均值 b 和扩展幅度 (Spread) s 是需要估计的参数。对于三角形分布,密度从 $b-s$ 开始取值为正,且在 $b-s$ 到 b 之间为线性增加,然后在 b 到 $b+s$ 之间为线性减小。曲线形状像帐篷或三角形。均值 b 和扩散幅度 s 是待估计的参数。均匀分布和三角形分布因为两边受限制,避免了正态分布和对数正态分布会导致部分决策者的系数大得不合理这一问题。也有学者使用 Rayleigh 分布,该分布在 $x=0$ 的其中一侧像对数正态分布一样,但是被研究证实其参数估计比对数正态分布容易。也有学者用截断正态分布。总之,学者们可以很自由地构造出符合某种行为特征的合理分布来表达密度函数。

决策者的可观测属性相关的口味的差异可以通过合理构造解释变量或混合分布来刻画。例如,成本除以收入,可以使得成本的价值或相对重要性随着收入的增加而降低。然后,“成本/收入”这一变量取随机系数,就可以表示相同收入的决策者中成本重要性的差异。成本的价值均值随收入增加而减少,但围绕均值的方差是固定的。此外,少量研究设定 $f(\beta)$ 时,把决策者的可观测属性纳入其中了,这意味着口味差异还依赖于决策者的属性。

2.3 误差-组成部分

混合 Logit 模型亦可简单地表达为:增加了用来表示不同选项的效用之间存在相关性的误差组成部分(error-components)。效用构造如下:

$$U_{nj} = \alpha' x_{nj} + \mu'_n z_{nj} + \epsilon_{nj} \quad (2.9)$$

式中 x_{nj} , z_{nj} ——与选项 j 有关的可观测变量组成的向量;

α ——固定系数组成的向量;

μ ——均值为零的随机项组成的向量;

ϵ_{nj} ——独立相同的极值分布。

式(2.9)中第二项是误差组成部分,它与 ϵ_{nj} 一起定义了选项效用的随机部分。也就是说效用中的随机(不可观测)部分是 $\eta_{nj} = \mu'_n z_{nj} + \epsilon_{nj}$,并且在 z_{nj} 的某构造条件下,在不同选项间是相关的。当 z_{nj} 都设为零时,就得到了标准 Logit 模型,于是也就有 IIA 特性以及受限性的取代状态。如果误差组成部分非零,那么选项的效用是相关的:

$$\text{Cov}(\eta_{ni}, \eta_{nj}) = E(\mu'_n z_{ni} + \epsilon_{ni})(\mu'_n z_{nj} + \epsilon_{nj}) = z'_{ni} W z_{nj} \quad (2.10)$$

式中, W 是 μ_n 的协方差。即便当误差组成部分是独立的,从而 W 是对角阵的时候(大多数研究者都如此构造),效用在各选项间亦为相关的。

各种相关形态(从而形成各种取代型式),可以通过合理选择进入误差组成部分的变



量来得到。例如,我们可以实现像嵌套 Logit 一样的模型。假如有 k 个不重叠的巢,巢中的选项的哑变量都为 1,否者为 0,误差组成部分构造为:

$$\mu'_n z_{nj} = \sum_{k=1}^K \mu_{nk} d_{jk} \quad \text{其中 } d_{jk} = \begin{cases} 1, & j \text{ 在巢 } k \text{ 中} \\ 0, & j \text{ 不属于巢 } k. \end{cases} \quad (2.11)$$

然后可以设定 μ_{nk} 服从独立相同的正态分布: $\mu_{nk} \sim \text{iid}N(0, \sigma_k)$, 于是随机项 μ_{nk} 仅仅进入巢 k 中的选项,不进入其他巢中的选项,从而只有同一巢中的选项间有相关性,而属于不同巢的选项之间没有相关性。 σ_k 表示相关性的 大小,有点像嵌套式 Logit 模型中的 “inclusive value coefficient” 的角色。

更加精确地讲,巢 k 中两个选项的协方差 $\text{Cov}(\eta_{ni}, \eta_{nj}) = E(\mu_k + \epsilon_{ni})(\mu_k + \epsilon_{nj}) = \sigma_k$, 巢 k 中每个选项的方差 $\text{Var}(\eta_{ni}) = E(\mu_k + \epsilon_{ni})^2 = \sigma_k + \pi^2/6$ (注: ϵ_{ni} 是极值分布,其方差是 $\pi^2/6$)。如果我们限定每个巢的方差都相同(即 $\sigma_k = \sigma, k = 1, 2, \dots, K$),就好比 we 限定嵌套 Logit 模型中所有巢中的内含效用系数都相同。这种限制确保了混合 Logit 模型在规模和水平上做了标准化。

允许不同巢的随机项取不同的方差就好比允许巢式 Logit 模型中各个巢的“内含重要性系数”(IVC)不同。另外,通过设定一些标示着重叠巢的哑变量,就可以刻画巢与巢之间有交叠的现象。最后,异方差 Logit 模型可以通过为每个选项设置误差组成部分来实现。专门有学者撰写研究报告来讲述如何合理构造变量,建立混合 Logit 模型。

误差组成部分和随机系数这两种构造方法实际上是相同的。在随机系数构造下,效用构造为 $U_{nj} = \beta'_n x_{nj} + \epsilon_{nj}$, 其中 β_n 是随机系数,系数 β_n 可以被分解为均值 α 和方差 μ_n 两部分,于是有 $U_{nj} = \alpha'_n x_{nj} + \mu'_n z_{nj} + \epsilon_{nj}$, 这时相当于定义 $z_{nj} = x_{nj}$ 。相反地,在误差组成部分构造形式如 $U_{nj} = \alpha'_n x_{nj} + \mu'_n z_{nj} + \epsilon_{nj}$, 这相当于认为对于 x_{nj} 这些变量的随机系数取为固定系数,而 z_{nj} 这些变量的随机系数取均值为零的随机变量,如果 x_{nj} 和 z_{nj} 重合了,那么这些变量的系数可以被认为是以均值 α 随机变化,其分布与 μ_n 分布形态相同。

虽然两种构造形式本质相同,但研究者思考和看待混合 Logit 模型的方式会影响具体的模型构造。例如,从随机参数考虑时,很自然地允许每个解释变量的系数是随机的并且甚至可以允许这些系数之间存在相关性。但是,我们的目标如果是通过使用误差组成部分来合理地表示取代型式,那么重点就是以精简的方式构造可以体现选项效用存在相关性的变量。值得一提的是,如果解释变量数量不多(例如 6 个),那么估计相应的系数的联合分布较易实现,建模结果往往可以达到预期效果。但是当变量很多(例如,20 多个)时,用随机参数构造形式往往不切实际,建模结果往往达不到预期的效果。最后要注意的关键是,不管是随机参数还是误差组成部分,混合分布都用来刻画未观测的因素的变化差异和相关性。而人们对这些不可观测的因素的理解和认识程度终归是有限的。

2.4 取代型式

混合 Logit 模型没有标准 Logit 模型具有 IIA 特性或限制性的取代型式,其选择概



率比 P_{ni}/P_{nj} 依赖于所有数据,包括选项 i 和 j 以外的其他选项的属性。其某一选项的选择概率变化百分比在给定另一选项的第 m 个属性的变化的条件下,计算如下:

$$\begin{aligned} E_{mix_{nj}}^m &= -\frac{1}{P_{ni}} \int \beta^m L_{ni}(\beta) L_{nj}(\beta) f(\beta) d\beta \\ &= -\int \beta^m L_{nj}(\beta) \left[\frac{L_{ni}(\beta)}{P_{ni}} \right] f(\beta) d\beta \end{aligned} \quad (2.12)$$

式中, β^m 是 β 的第 m 个元素。以上所示的弹性计算式对于每个选项 i 是不同的。一个 10% 的变化在某个选项上发生了,并不意味着在另一个选项上也发生。实际情况则是:取代型依赖于根据经验所确定的变量的具体构造形式和选用的混合分布。

要注意的是:选项概率的变化百分比依赖于在不同的 β 取值上 $L_{ni}(\beta)$ 和 $L_{nj}(\beta)$ 的相关性,而这种相关性则是由研究者的变量构造型式和混合分布构造型式决定的。比如说,若研究者要表示“选项 j 的改善对于选项 i 的促进作用比在选项 k 上的更显著”这一情形,就可以构造一个分量,该分量使得选项 i 和 j 之间存在正相关性,而使得选项 k 和 j 之间不存在相关性或存在负相关,然后再配上一个允许变量系数变化的混合分布函数,就可以了。

2.5 对任何随机效用模型的近似

McFadden, Train(2000)表明,只要合理地选择解释变量和混合分布,混合 Logit 模型可以非常准确地近似几乎任何一种随机效用模型(RUM)。下面给出了直觉上的解释。假设真实模型是 $U_{nj} = \alpha'_n z_{nj}$, 其中 z_{nj} 是选项 j 相关的变量, α 则可以服从任何一种分布 $f(\alpha)$ 。任何随机效用模型都可以表达为此形式。如果令 $z'_{nj} = \langle x'_{nj}, d_j \rangle$ 和 $\alpha' = \langle \beta'_n, \epsilon_{nj} \rangle$ 以及 $f(\alpha)$ 为 β_n 和 ϵ_{nj} 的联合密度函数,那么就可以得到更加常规的表达式是 $U_{nj} = \beta'_n x_{nj} + \epsilon_{nj}$, 在确定 α 的条件下就可以确定某个决策者的选择,因为 U_{nj} 对于每个选项 j 来说是已知的。因此条件概率为:

$$q_{ni}(\alpha) = I(\alpha'_n z_{ni} > \alpha'_n z_{nj}, \forall j \neq i) \quad (2.13)$$

式中, $I(\cdot)$ 是括号内的事件是否发生的指标变量。这一条件概率或者等于 0, 或者等于 1, 因此,在这种意义下是确定性的,即在知道了随机项的条件下,决策者的选择可以被确定下来。非条件概率是 $q_{ni}(\alpha)$ 在 α 取值范围里的积分结果:

$$Q_{ni} = \int I(\alpha'_n z_{ni} > \alpha'_n z_{nj}, \forall j \neq i) f(\alpha) d\alpha \quad (2.14)$$

我们可以用混合 Logit 模型来近似以上概率。通过变化 λ 来变化效用的规模,即 $U_{nj}^* = (\alpha/\lambda) z_{nj}$ 。这种规模变化不会改变模型,因为决策行为不受规模变化的影响。然后,我们加入一个 iid 极值分布项: ϵ_{nj} 。这一添加项倒是会改变模型,因为它影响每一个选项 j 。我们添加此项,是因为这样做可以得到一个混合 Logit 模型。添加此项是无害的。基于以上效用设定情况,混合 Logit 概率为: