

中国高铁出版工程——运营管理系列

高速铁路 客流时变需求预测研究

魏堂建◎著



西南交通大学出版社

图书在版编目 (C I P) 数据

高速铁路客流时变需求预测研究 / 魏堂建著. —成都: 西南交通大学出版社, 2020.12
ISBN 978-7-5643-7881-3

I. ①高… II. ①魏… III. ①高速铁路—客流—运输需求—预测—研究 IV. ①U293.13

中国版本图书馆 CIP 数据核字 (2020) 第 244577 号

Gaosu Tielu Keliu Shibian Xuqiu Yuce Yanjiu

高速铁路客流时变需求预测研究

魏堂建 / 著

责任编辑 / 穆 丰

助理编辑 / 宋浩田

封面设计 / 曹天擎

西南交通大学出版社出版发行

(四川省成都市金牛区二环路北一段 111 号西南交通大学创新大厦 21 楼 610031)

发行部电话: 028-87600564 028-87600533

网址: <http://www.xnjdcbs.com>

印刷: 四川煤田地质制图印刷厂

成品尺寸 170 mm × 230 mm

印张 10.25 字数 230 千

版次 2020 年 12 月第 1 版 印次 2020 年 12 月第 1 次

书号 ISBN 978-7-5643-7881-3

定价 68.00 元

图书如有印装质量问题 本社负责退换

版权所有 盗版必究 举报电话: 028-87600562



魏堂建，博士，华东交通大学交通运输与物流学院交通运输系副主任，世界交通运输大会（WTC）轨道交通-铁路旅客运输-铁路客运组织委员。

工作期间荣获江西省教学成果二等奖、江西省教育科学优秀成果二等奖、江西省高等学校微课比赛二等奖、江西省精品在线开放课程《铁路行车组织》主讲人、华东交通大学首届“天佑主讲教师”、华东交通大学“优秀教师”和“最受学生喜爱教师”等荣誉。

在国内外权威期刊和会议发表学术论文12篇，申请或授权国家发明专利3项、软件著作权1项，主持江西省教育厅科学技术研究项目重点项目，江西省高校人文社科基金，参与国家自然科学基金高铁联合基金、国家自然科学基金面上项目、青年项目和地区项目等项目，主持和参与编写教材4本。

主要研究方向：

1. 高速铁路客流大数据分析
2. 高铁客流时变需求估计
3. 列车开行方案优化和运行图编制
4. 交通流建模仿真与分析

文轨车书 交通天下

<http://www.xnjdcbs.com>

策划编辑 / 黄庆斌 黄淑文 周 杨

责任编辑 / 穆 丰

助理编辑 / 宋浩田

封面设计 / 曹天擎



交大e出版
微信购书|数字资源



官方天猫店
上天猫 买正版

前 言

旅客出行需求的相关信息是编制列车开行方案、优化票额分配和制订售票计划等运输组织优化和收益管理的基本输入数据。随着高铁网络化运营进程的推进，铁路的运输能力得到了极大提升，旅客对于高铁所提供的运输服务也有了进一步的要求：在出行需求总量得到满足的前提下，同时希望自身的出行需求得到及时响应，即满足高铁旅客的时变需求。于是，精准预测高铁客流时变需求是编制高质量的运输计划和制定科学管理决策，进而保证高铁服务质量和提高运营效率的基本前提。

本书针对如何获取高铁客流时变需求这一问题进行理论研究。根据高铁客流时变需求的特点，将其分解为时变需求分布估计和日客流量预测两个子问题。针对时变需求分布估计问题，分别设计了基于客流量的高铁旅客时变需求分布逆分配估计法和基于购票记录的最大熵估计模型。针对日客流量预测问题，设计了适用于中期预测的双层平行小波神经网络模型。本书编写过程中的主要工作和本书的创新点如下：

(1) 对高铁客流时变需求预测问题进行了剖析。提出获取高铁客流时变需求需要解决的核心问题是时变需求分布估计和高铁日客流量预测；其中，前者需要解决的核心问题是如何消除列车运行图、列车能力、票价费用等因素对高铁旅客出行选择的影响，以获得一天内高铁旅客在运营时间段的出行需求分布；而后者需要解决的核心问题是如何保证在进行中期日客流量预测（120天左右）时预测精度。

(2) 设计了基于客流量的时变需求分布逆分配估计方法。模拟高铁旅客购票选择过程，以任意有效乘车方案能力利用饱和或车票售罄（简称能力饱和）为分界点，将O-D间的整个购票过程划分为若干购票阶段。提出了优于关系链模型，体现各个乘车方案之间的能力饱和顺序；然后

结合最长优于关系链，提出了两种典型的乘车方案饱和情况，并基于这两种典型情况分别设计了单购票阶段逆向分配估计法和多购票阶段逆向分配估计法，通过算例测试和参数灵敏度分析验证了方法的有效性。

(3) 建立了基于购票操作数据的时变需求分布最大熵估计模型。从铁路售票系统中提取信息来构建旅客购票操作链和预售期各时刻的可用乘车方案集，结合这两者信息，并利用屋顶模型设计了推算旅客期望出行时间范围的方法，然后，基于各旅客的期望出行时间范围，建立了时变需求最大熵估计模型并设计了对应的求解算法，并且通过算例一展示了算法计算详细过程，再通过算例二检验了最大熵估计算法的精度和有效性。

(4) 提出了高铁旅客日客流量中期预测方法。在对历史数据特征进行分析的基础上，构建了高铁旅客日客流量预测的 DLP-WNN 双层平行小波神经网络模型。其原理是通过综合两个平行的小波神经网络的输出结果来获得预测输出，该方法既可以延续近期日客流量的总体趋势，又体现了不同日期（特别是节假日）之间的需求量的差异，以此保证进行中期预测（120 天左右）的精度，并且通过京沪高铁中四种不同出行距离的典型 O-D 进行日客流量预测，验证和评价了所提出的预测模型。

全书由魏堂建撰写，研究生韩忠贵、秦昱和陶乐风负责了部分图表的编辑工作，本科生黄生辉、刘妍、苗菁菁对参考文献和附录进行了更新和整理。中南大学博士生导师史峰教授对本书的出版提出了宝贵的指导意见，在此表示衷心感谢；同时感谢中南大学李夏苗教授、符卓教授、王璞教授、邓连波教授、秦进教授和徐光明副教授积极推动本书的出版；感谢华东交通大学交通运输与物流学院的领导及同事在作者编写本书过程中提供的帮助和支持。

鉴于作者水平有限，书中难免存在疏漏之处，敬请各位同行专家学者批评指正，以使该书得到不断完善，为高速铁路的发展尽献绵薄之力。

魏堂建

2020 年 10 月于华东交通大学孔目湖畔

目 录

第 1 章 绪 论	1
1.1 研究背景	1
1.2 研究意义	2
1.3 研究思路	5
1.4 研究内容	6
1.5 小 结	8
第 2 章 国内外文献综述	9
2.1 时变需求分布估计国内外研究	9
2.2 旅客日客流量预测国内外研究	16
2.3 小 结	18
第 3 章 高速铁路旅客时变需求预测问题分析	19
3.1 时变需求分布估计问题分析	19
3.2 日客流量预测问题分析	24
3.3 小 结	24
第 4 章 旅客购票选择过程及客流分配屋顶模型	25
4.1 屋顶模型和有效乘车方案集	25
4.2 有效乘车方案吸流区间	28
4.3 旅客购票选择过程及购票阶段划分	29
4.4 小 结	30
第 5 章 时变需求分布逆向分配估计法	31
5.1 逆分配估计法原理	31
5.2 单购票阶段逆向分配估计法	38

5.3	多购票阶段逆向分配估计法	49
5.4	有效性检验	59
5.5	小 结	60
第 6 章	时变需求分布最大熵估计法	62
6.1	当前可用乘车方案集及购票操作链	62
6.2	旅客期望出行时间范围推算	65
6.3	时变需求分布最大熵估计	74
6.4	算例分析一	80
6.5	算例分析二	90
6.6	小 结	99
第 7 章	高速铁路日客流量中期预测	100
7.1	旅客日客流量特征提取	100
7.2	预测模型及方法	108
7.3	算例分析	115
7.4	小 结	125
第 8 章	结论与展望	126
参考文献		129
附 录		139
附录 1	BJ—CZ 各种预测方法结果对比表	139
附录 2	BJ—QF 各种预测方法结果对比表	144
附录 3	BJ—NJ 各种预测方法结果对比表	149
附录 4	BJ—SH 各种预测方法结果对比表	154

第1章 绪论

1.1 研究背景

铁路运输是中国综合运输体系中的骨干成员和主要运输方式之一，其在推进社会经济发展中起到至关重要的作用。高速铁路是在普通铁路基础上的进步和突破，“高速”是一个相对的概念，目前国际铁路联盟（International Union of Railways, UIC）关于高速铁路的定义为：新建线路设计速度达到 250 km/h 及以上，或者既有线经过升级改造（直线化、轨距标准化）的设计速度达到或者超过 200 km/h 的铁路^[1]属于高速铁路。中国国家铁路局关于高速铁路的定义为：列车开行速度在 250 km/h 以上的新建铁路线，以及初期运营速度不小于 200 km/h 的铁路客运专线都属于高速铁路^[2]。

进入 21 世纪后，我国的高速铁路发展迅速。自 2002 年建成第一条秦沈客运专线以来，又相继于 2008 年建设运营了 350 km/h 的京津城际、250 km/h 的合宁高铁以及于 2009 年建设运营的 350 km/h 的武广高铁等。目前，中国的“四纵四横”高速铁路网络已经基本建设完成并投入运营。2018 年，中国铁路全年完成旅客运输 33.7 亿人次，比上年增长 9.4%，完成旅客运输周转量 14 146.6 亿人公里，比上年增长 5.1%^[3]。截至 2018 年年底，中国铁路的营业里程达到了 13.1 万公里，其中高速铁路营业里程为 2.9 万公里以上，位居世界第一，且超过世界高速铁路总里程的三分之二^[4]。

在过去十余年间，中国铁路改革发展成效显著，基础设施建设持续加快，运输能力大幅提升，服务水平明显提高。同时，为了满足快速增长的旅客运输需求，2016 年中国发布的《中长期铁路网规划》^[5]和 2017

年发布的《铁路“十三五”发展规划》^[6]中都指出,我国将在目前“四纵四横”高速铁路骨架网络的基础上,规划构建“八纵八横”高速铁路主通道网络,如图 1-1 所示。2020 年,中国的铁路网规模将达到 15 万公里,其中高速铁路 3 万公里以上,覆盖 80% 以上的大城市;到 2025 年,铁路网规模达到 17.5 万公里左右,其中高速铁路达到 3.8 万公里左右^[3]。



图 1-1 中国高速铁路“八纵八横”规划图(请扫码观看原图)

上述规划建成的现代高速铁路网将连接中国的主要城市群,基本连接省会城市和其他 50 万以上人口的大中城市,形成以特大城市为中心覆盖全国、以省会城市为支点覆盖周边的高速铁路网。同时,为了让旅客的出行更为便捷,预计将实现动车组列车承担旅客运量占比达到 65%;北京至大部分省会城市实现 2~8 h 通达、相邻大中城市间 1~4 h 通达的高速铁路交通圈和主要城市群内 0.5~2 h 通达的高铁交通圈^[5]。高速铁路将为旅客提供更加安全可靠、优质高效和舒适便捷的运输服务。

1.2 研究意义

高速铁路与其他运输方式相对比较,具有运行速度快、旅行总时间短、安全性能高、运送旅客能力大、节能减排、对环境污染小、全天候运行、受气候影响因素小、正点率高、乘坐舒适、经济效益高等特点^[7-8],随着我国高速铁路网络的形成,网络化、高速度、高密度等全新运营模式正在改变人们的出行方式。

对于铁路客运部门来说,制定运输计划和进行日常的铁路运输生产

活动的主要目标是尽可能地满足旅客出行需求。在传统普速铁路线路上,由于列车速度低并且旅客列车和货物列车混跑以至于线路能力紧张,导致开行的旅客列车数量受到限制,使其在列车开行方案制定和运行图编制时,都是以尽可能地满足铁路网络上各个 O-D 站点间每天的旅客出行需求量为目标^[9-12]。然而,高速铁路由于列车速度得到提高,线路的运输能力得到极大提升,于是高铁线路上就可以开行更多的列车,例如,北京与上海这两个城市之间的距离约为 1 300 km,在 2006 年两者之间还未开通高铁列车时,北京至上海的普通铁路的服务频率是每天约 10 趟列车;而随着高速铁路的建成并投入运营,比如在 2019 年,北京和上海这两个城市之间的高铁服务频率是每天约 42 趟列车(通过网址 <https://www.12306.cn> 查得)。随着高铁线路运输能力的提升,高铁 O-D 间输送旅客的能力相对普通铁路来说提高了很多,原本铁路运输能力不足的状况几乎很少再出现(个别节假日除外),渐渐地,整个社会对于高铁所提供的旅客运输服务又有了新的要求,不再仅仅关注于每天的出行需求量是否得到满足,更加强调对能否便捷地出行等的关注。例如,是否在自己期望的出发时间点有列车开行从而不用给自己带来过多的等待时间或者被迫进行出发时间的调整。于是,越来越多的高速铁路研究学者^[13-18]提出在编制高铁列车运行计划时,要尽可能地满足高速铁路客流的时变需求,以最大限度地提升高铁服务质量。

高速铁路客流时变需求包括以下两个方面的内容:

- (1) 高铁 O-D 间一天内的旅客时变需求分布;
- (2) 高铁 O-D 间时变的日客流量(以下简称时变日客流量)。

在上述两方面内容中,高速铁路 O-D 间一天内的旅客时变需求分布(以下简称时变需求分布)是指:在高铁 O-D 间,一天运营时段内不同时间点的旅客出行需求量是不同的,于是旅客出行需求量关于该天运营时间的分布曲线就是该 O-D 的时变需求分布;而高速铁路 O-D 间时变的日客流量是指该 O-D 间的客流量会随着日期的变化而波动。例如在不同的工作日、周末或者节假日,O-D 间的日客流量是存在波动差异的。另外,更加广义的时变需求应该还包括以月或者年

为观测角度的客流需求量变化情况，它们通常是高铁线网规划（例如“铁路‘十三五’规划”或“中长期铁路网规划”）时的基础输入数据，而本书所考虑的旅客时变需求主要是针对高铁运营层面的，所以仅仅考虑一天内的时变需求分布和一段时期内不同日之间的时变客流量就足够。

近年来，为了更加适应旅客运输市场，我国高铁实行了“动态”的列车开行方案，采用“按流开车”和“一日一图”的运输组织形式，即针对高铁旅客时变需求，采用动态的运输计划来提升运输服务质量。从高速铁路运输供需市场的角度进行分析可以知道，运营管理部门可以根据 O-D 间的日客流量变化情况确定动态的高铁列车日开行频率/数量，并根据 O-D 间的时变需求分布来确定该 O-D 间一天内高铁列车的发车间隔/时刻，以此来尽可能地满足旅客的出行需求，提升高速铁路的服务水平。

在上述“按流开车”“一日一图”运输组织模式和满足时变需求的列车运行计划编制过程中，高铁客流时变需求是作为输入数据被使用的，因此如何获取高铁客流的时变需求就变成了一个关键的基础性问题。

传统的获得旅客出行需求数据的方式通常是采用 O-D 调查或者应用需求量预测的相关方法来完成，但是调查方法工作量巨大、持续时间长、成本高、数据的有效时间短、难于多次重复并且精度不一定有保证，显然不能用在时变需求预测上面。同时，当前的需求量预测方法绝大多数都是短期预测技术，这个短期通常是指输入数据接下来的 1 期、1 天或者 1 步，但这种方法无法直接用于列车运行计划编制时的需求预测。因为当前我国车票预售期是 30 天，即从当天开始之后 30 天的运输计划都已经确定且提供了车票供旅客购买，于是在编制新的列车开行方案和编制列车运行图等运输计划时，就至少需要考虑 30 天之后的旅客需求情况。另外，一套列车运行计划的使用“服役”时间至少有 2~3 个月，显然短期预测技术无法有效地对未来 120 天这种时间长度的日客流量进行准确的预测。另外，对于 O-D 间的时变需求分布，当前几乎没有可以直接利用的方法来进行有效估计。

于是，开展如何获取高铁客流的时变需求这个问题的研究就显得尤为必要。

本书致力于高铁客流时变需求预测理论研究。首先，将研究问题拆分为高速铁路 O-D 间的时变需求分布估计和日客流需求量预测这两个子问题；其次，针对时变需求分布估计问题，分别提出了逆向分配估计法和最大熵估计法；最后，对于 O-D 间日客流量预测问题，提出了适合于中期（120 天）预测的双层平行小波神经网络预测法。通过上述相关时变需求预测方法的提出，为高速铁路的运输计划编制和决策管理提供更加精准的数据支持。

1.3 研究思路

高速铁路客流时变需求数据是编制高铁运输计划，进行日常运输组织管理的重要输入数据，时变需求预测数据的质量会直接影响到铁路运输计划的编制效果，进而影响铁路服务水平质量和运营效益。我们大体按照如下思路进行高铁客流时变需求预测研究。

第一，根据高速铁路旅客时变需求的内涵，将问题分解为高铁旅客时变需求分布估计和日客流量预测这两个子问题，针对这两个问题各自的特征，分别设计相应的求解方法。

第二，铁路售票系统为进行高铁旅客时变需求预测和估计提供了重要的数据来源。铁路售票系统不仅记录了旅客选择高铁列车时的原始购票操作信息，同时还记录了所有旅客对于各个高铁列车的选择结果，即各个列车的客流量，通过对这些信息进行提取，可以为我们的问题分析提供重要的基础数据和信息。

第三，通过推算各个旅客的期望出行时间范围，来估计高铁客流的时变需求分布。在进行时变需求分布估计时，并不是去确定每一个旅客的精确期望出行时间点，而是通过各个旅客所选择的出行列车，去逆向推算他们各自的出行时间范围，在此基础上去估计时变需求分布。

第四，分别采用模拟方式和提取旅客购票操作信息的方式来分别得到 O-D 间所有高速铁路旅客对于各个列车的购票选择顺序，据此推断出依次购票的各个旅客的期望出行时间范围，然后结合逆分配原理和最大熵模型去进行时变需求分布估计。

第五，高速铁路旅客日客流量预测方法需要具备进行中期（120 天）预测操作的能力。将预测模型设计为双层平行神经网络，第一层网络主要体现近期若干天的日客流量对当前的日客流量影响，第二层网络体现预测日当天的时间属性和节假日属性对于该天日客流量的影响，综合两个网络的输出以预测当前的日客流需求量。由于第二层网络中预测期每天的时间属性信息和节假日属性信息都可以提前确定，从而保证了进行中期日客流需求量预测的精度。

1.4 研究内容

本书研究高速铁路客流时变需求预测方法。本书首先讨论和分析时变需求预测这个问题的核心内涵，将其分解为时变需求分布估计和日客流量预测两个子问题分别进行研究；其次针对高铁旅客时变需求分布估计问题，分析基于屋顶模型的旅客购票选择过程；再次，分别采用模拟旅客购票的思路和利用历史购票数据信息的思路，设计时变需求分布逆向分配估计法和最大熵估计法；最后，针对高铁旅客日客流量预测问题，根据历史数据，获得高铁日出行需求变化规律和特点，并构建适合中期预测的日客流需求量预测模型。

本书的研究内容结构如图 1-2 所示，具体的研究内容如下：

第 1 章为绪论。该章首先分析了本书的研究背景以及研究意义，然后介绍了书的主要研究思路和研究内容。

第 2 章为国内外文献综述，分别对旅客日客流量预测和时变需求分布估计这两个方面进行国内外论文研究综述分析。

第 3 章为时变需求预测和估计问题分析。该章首先将研究内容分解为高铁日客流量预测问题和高铁旅客时变需求分布估计问题；然后，针对时变需求分布估计问题，其需要解决的核心问题是如何消除

列车运行图、列车能力、票价费用等因素对高铁旅客出行选择的影响,以获得旅客的出行需求分布;最后对于日客流量预测问题,提出其需要解决的核心问题——如何在连续多天的日客流量预测时保证预测精度。

第4章为旅客购票选择过程及客流分配屋顶模型。这一章主要介绍屋顶模型的原理以及通过该模型所能够得到的有效乘车方案及其对应的吸流区间。通过屋顶模型可以刻画高铁旅客的出行购票选择过程,为后续的时变需求分布估计奠定基础。

第5章为时变需求分布的逆向分配估计法。该章首先介绍逆向分配估计法的原理及基本操作步骤;然后提出优于关系链来体现乘车方案之间的能力饱和顺序,并基于此提出两种典型的乘车方案能力饱和情况;其次,各结合一种典型情况分别设计单购票阶段逆向分配估计法和多购票阶段逆向分配估计法,并各自进行算例分析和灵敏度分析;最后,在该章末尾还会对上述两种方法进行有效性检验。

第6章为时变需求分布的最大熵估计法。该章从铁路售票系统中提取的旅客购票操作信息来进行高铁旅客时变需求分布估计。首先将每个旅客的购票操作行为构建成该旅客的购票操作链,并提出基于屋顶模型的旅客期望出行时间范围推算算法;其次,基于各个旅客的期望出行时间范围,建立时变需求最大熵估计模型并设计对应的求解算法;再次,通过算例一展示求解旅客期望出行时间范围和估计时变需求分布的具体计算过程和结果;最后再通过算例二来展示所提出的算法的估计精度。

第7章为高速铁路日客流中期预测。这一章首先基于历史数据,对高铁旅客的日客流量进行特征分析,并获得影响日客流量的日期属性和节假日属性;其次构建日客流量预测的双层平行小波神经网络模型,模型以预测型的近期日客流量数据为子网络1的输入;最后以确定型的预测期各天的时间属性值和节假日属性值为子网络2的输入,综合两个子网络以获得整个预测输出,就既可以延续日客流量趋势又体现不同日期之间的需求量差异,以此保证预测精度。

最后的内容为结论及展望。这一部分主要是归纳和总结本书的主要工作和创新点,并对未来的相关扩展研究做介绍。

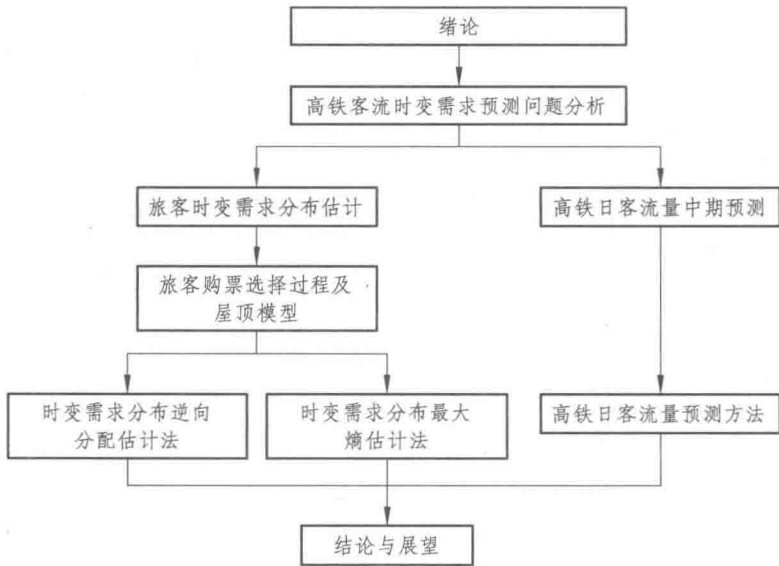


图 1-2 本书内容结构安排

1.5 小结

本章从研究背景入手，介绍了当前我国的高铁运输情况，并通过研究意义这一节引入了高速铁路旅客时变需求的两个方面的内涵，即高速铁路 O-D 间一天内的旅客时变需求分布和高铁 O-D 间时变的日客流量，然后在后两节介绍了本文的主要研究思路、研究内容和研究结构等。

第2章 国内外文献综述

通过第1章的分析可以知道,高速铁路旅客的时变需求预测包括O-D间时变需求分布估计和日客流量预测两个方面的内容。下面分别就这两个问题的国内外研究现状进行分析。

2.1 时变需求分布估计国内外研究

目前国内外学者对于高速铁路旅客在一天内的时变需求分布估计的研究涉及较少,而在过去的几十年,国内外相关学者对于交通领域时变需求估计或预测的研究主要集中在航空动态需求预测、道路网络动态O-D估计/反推、城市轨道交通和公交网络时变O-D估计这几个方面,下面分别围绕这几个方面进行介绍和分析。

1. 航空动态需求预测

在航空运输领域,运输市场中多家航空公司之间存在相互竞争,旅客为了使得自身的出行成本最低会在不同的航空公司之间进行出行选择,于是各个航空公司根据自身特点针对运输市场进行收益管理。收益管理的目标就是“将合适的席位以合适的价格出售给合适的消费者”^[19],而在进行收益管理时,精确的旅客需求预测是其分析决策的核心输入内容^[20]。

航空领域对于动态需求的预测主要基于历史购票数据。航空购票数据分析的一个核心任务就是利用受约束的历史购票数据(censored historical booking data)来估计不同价格类型下的旅客需求量^[21]。历史购票数据是受约束的数据,是因为在旅客购票过程中,一旦达到了一个系统设定的购票限制,后续旅客再次购票将会被拒绝,而这些拒绝的需求