

DASHUJU YU CHENGSHI JINGJI
FAZHAN YANJIU

大数据与城市经济 发展研究

肖良武 孙庆刚 张攀春 黄 臻 编著



北京邮电大学出版社
www.buptpress.com

大数据与城市经济发展研究

肖良武 孙庆刚 张攀春 黄 臻 编著

北京邮电大学出版社

• 北京 •

内 容 简 介

本书内容聚焦大数据时代出现的经济、管理新现象和新问题,关注大数据战略驱动下贵阳城市经济和区域经济发展的新情况。内容具体分为三部分:大数据篇、城市经济篇和区域经济篇。大数据篇重点关注了大数据对精准营销、电商模式、融资模式和物流服务等问题;城市经济篇重点关注了大数据时代贵阳市生态补偿、创新驱动、产业结构升级和智慧旅游等问题;区域经济篇关注了农业现代化、绿色金融和贵州省两高经济带发展问题。

图书在版编目(CIP)数据

大数据与城市经济发展研究 / 肖良武等编著. -- 北京:北京邮电大学出版社, 2019.5

ISBN 978-7-5635-5726-4

I. ①大… II. ①肖… III. ①数据处理—应用—城市经济—经济发展—研究—贵阳

IV. ①F299.277.31-39

中国版本图书馆 CIP 数据核字(2019)第 090662 号

书 名 大数据与城市经济发展研究

编 者 肖良武 孙庆刚 张攀春 黄 臻

责任编辑 沙一飞

出版发行 北京邮电大学出版社

社 址 北京市海淀区西土城路 10 号(100876)

电话传真 010-82333010 62282185(发行部) 010-82333009 62283578(传真)

网 址 www.buptpress3.com

电子信箱 ctrd@buptpress.com

经 销 各地新华书店

印 刷 北京建宏印刷有限公司

开 本 787 mm×960 mm 1/16

印 张 14.5

字 数 276 千字

版 次 2019 年 5 月第 1 版 2019 年 5 月第 1 次印刷

ISBN 978-7-5635-5726-4

定价: 78.00 元

如有质量问题请与发行部联系

版权所有 侵权必究

前 言

大数据开启了一个新时代,新的商业模式、新的经济形态和新的生活方式正在形成并不断演进。贵州省作为西部内陆省份,按照“守底线、走新路、打造升级版”的内在要求,依靠创新驱动,实施大数据战略行动,最终实现后发赶超、跨越发展的目标。贵阳市作为贵州省的中心城市,引领着全省大数据产业发展,在省域范围内对经济增长发挥着重要的作用。近年来,贵阳市委、市政府结合本市实际情况,认真贯彻落实中央、省有关创新发展精神,主动适应新常态,引领新常态,不断推进改革开放,初步形成了以大数据引领科技创新、金融创新为支撑,以政策创新为抓手,以制度创新为突破口的创新发展格局。

贵州省贵阳市依靠国家支持政策,不断优化大数据产业发展环境。中央及国家相关部委出台了一系列促进贵州省贵阳市发展的政策,将贵阳市列入全国首批国家创新型试点城市等。贵州省人民政府印发了《关于加快大数据产业发展应用若干政策的意见》《贵州省大数据产业发展应用规划纲要(2014—2020年)》等发展大数据的意见及规划,贵阳市委、市政府也先后出台了《关于加快发展高新技术产业和现代制造业的意见》《贵阳大数据产业行动计划》《关于加快发展贵阳大数据产业的实施意见》等政策措施,从财税、金融、人才、土地等各个方面,为支持大数据产业发展提供了政策保障,也为本地实现大数据全产业链集聚提供了重要推动力。

2017年,贵阳市委、市政府明确提出:到2020年,贵阳市将基本建成全域块数据城市、具有影响力的“中国数谷”,完成国家大数据试验示范任务,形成一批大数据示范应用,大数据产业竞争力进入全国城市第一梯队,建成国家大数据产业发展集聚区。事实上,贵阳市大数据全产业链项目布局正在不断推进。贵阳市以大数据公共平台建设为支撑,以“块数据”城市应用示范为抓手,以政府数据开放和市场

数据交易为目标,围绕“1+N”的产业布局,打造大数据内容中心、服务中心和金融中心,已形成“一城两带六核”的中关村贵阳科技园总体布局。

大数据为贵阳提供了前所未有的后发赶超的机遇,促进了城市产业结构优化和城市品位与城市竞争力提升,推动了区域经济快速发展。在此背景下,本书受贵州省哲学社会科学创新团队项目、贵阳市科技局贵阳学院专项资金项目资助,将内容聚焦于大数据时代经济社会的新现象、新问题,关注大数据战略驱动下贵阳市城市经济和贵州区域经济发展的新情况。内容具体分为三部分:大数据篇、城市经济篇和区域经济篇。大数据篇重点关注大数据对精准营销、电商模式、融资模式和物流服务等问题;城市经济篇重点关注大数据时代贵阳市生态补偿、创新驱动、产业结构升级、智慧旅游及城市等级体系等问题;区域经济篇重点关注农业现代化、绿色金融和贵州省两高经济带发展问题。

编者

2019年2月于贵阳学院贵州城市经济研究所

目 录

第一篇 大数据

大数据实名需求背景下的收益管理网络无约束估计方法	2
大数据对高校统计学科发展的影响分析	12
大数据时代企业精准营销策略分析	20
论数据挖掘在精准营销中的应用	27
大数据时代贵阳市旅游目的地营销策略研究	34
大数据时代网络环境下个人信息的安全保护	41
大数据时代电子商务模式创新研究	50
Hadoop 大数据中心和可视化管理结构分析与设计	56
大数据金融下小微企业的融资模式创新研究	68
大数据对贵阳产业结构的影响研究	76
大数据时代餐饮业 O2O 营销模式研究	84
大数据时代下物流服务的问题与策略研究	91

第二篇 城市经济

黔中经济区中心城市考察与城市等级体系研究	98
贵阳市推动创新驱动与跨区域合作研究	107
贵阳完善资源有偿使用和生态补偿制度研究	127
金融共生视角下贵阳市社区银行发展路径研究	138
新常态下智慧旅游驱动贵阳城市建设研究	148
双龙航空港经济区发展战略研究	157
新常态下促进贵阳科技服务业发展的对策研究	168
贵阳市加快推进服务外包发展研究	176
新常态下贵阳经济结构转型升级的路径研究	184
贵安新区建设对贵阳市经济的影响研究	192

第三篇 区域经济

贵州省两高经济带发展研究 199

绿色金融：中央政府与地方政府的投资博弈 208

资源禀赋与农业现代化路径选择：来自国外的经验借鉴 218

學問戰國編年表

馬建忠著

第一篇

大数据

大数据实名需求背景下的收益管理网络 无约束估计方法

郭 鹏^①

一、引言

由于在收益管理系统(Revenue Management System, RMS)应用过程中需要进行准确的无约束需求预测(Unconstrained Demand Forecasting),对乘客“初始需求(Primary Demand)”的无约束估计成了现代航空客运业亟待解决的难点问题之一^[1-2]。“初始需求”是指乘客首次发出购买请求时的机票需求,它对应于具体的机票类型、起飞时间等^[3]。在收益管理系统的存量控制和动态定价策略实施过程中,乘客“初始需求”通常受到舱位存量可用性和价格的影响,导致了乘客可能无法购买到理想中的机票,并产生了乘客需求的“再现(Recapture)”和“溢出(Spill)”效应^[4];前者是乘客转移购买公司其他舱位机票的“再现需求(Recaptured Demand)”,在航空客运业中,需求再现率达到了15%~55%^[5];后者是乘客放弃购买或转移购买其他公司机票的“溢出需求(Spilled Demand)”。这两种效应受短视型和策略型乘客选择行为的共同作用,并造成了预订系统中历史预订数据的不完备性等问题^[6]。

Guo^[1]等和 Azadeh^[2]等的研究表明,收益管理中的需求无约束估计方法按照所采用的技术手段可分为基础法、数理统计法和选择模型法。其中,基础法对历史预订数据采取简单的非参数处理方式,主要包括忽略法、舍弃法、替代法和直接观察法。若在收益管理实践中忽略需求的“再现”和“溢出”效应,比如,将预订系统中所记录到的乘客需求数据不作处理(忽略法),或直接舍弃存在需求“溢出”效应的数据(舍弃法),并以相似产品的不存在“溢出”效应的需求数据替代(替代法),如果将它们直接用于预测,不仅会导致预测模型对未来乘客“初始需求”的低估^[7]或高估问题^[3-4],还会影响公司的收入,并给收益管理系统的开发和实施过程带来困难。

为了提高收益管理系统无约束需求预测的准确性,在预订系统仅能获取“受截尾需求(Censored Demand)”数据(仅由购买请求得到满足的历史乘客“初始需求”和“再现需求”组成的可观察历史预订数据)的情况下,有必要利用数理统计法和选

^① 作者简介:郭鹏(1984—),男,四川宜宾人,贵阳学院副教授,博士,研究方向:电子商务。

择模型法对乘客真实的“初始需求”进行系统的无约束估计计算。它们通常假设无约束需求是随机的,并且符合一种或多种概率分布,这两种类型的方法均包括对概率分布参数的估计^[8]。

近年来,虽然对需求无约束估计方法的研究取得了较大进展,并且部分方法随着商业收益管理软件的发行而获得了成功应用,但现有的方法仅针对单航段的情况,并未考虑收益管理实践中产品间的网络替代效应。因此,有必要对网络无约束估计方法进行研究,以提高收益管理系统在网络环境下的预测精度。同时,随着大数据时代的来临、近年来信息技术的不断进步,以及航空客运业机票预订机制的成熟和实名制政策的实施^[9],航空客运公司机票预订平台直接获取乘客需求数据的能力得到大幅提升,在收益管理网络(Origin and Destination, O&D)市场环境中,不仅能够观察到实名乘客的“受截尾需求”数据,还能直接记录到实名乘客的部分相关“溢出需求”,基于乘客实名记录(Passenger Name Records, PNR)对真实“初始需求”进行预测也已成为了可能^[10-11]。

为此,本文首先扩展了现有无约束估计方法仅适用于单航段收益管理实践的前提假设,在考虑产品间的网络替代效应基础上,利用数据挖掘思想,提出了基于乘客实名记录的“初始需求”分解查找法;在此基础上,针对可能存在的取消预订和 no-shows 行为,建立了乘客历史实名实际“初始需求”的无约束估计计算模型;然后,通过数值算例说明了所提网络无约束估计方法的可行性;最后,在结论部分对所提方法的应用前景进行了探讨。

二、基于乘客实名记录的网络无约束估计方法

1. 符号及变量定义

结合航空客运收益管理实践,在考虑网络替代效应的情况下,本文基于乘客实名记录的“初始需求”分解查找法以及实际“初始需求”无约束估计计算模型的符号及变量定义如下:

- (1) M : 航线网络中的出发地集合;
- (2) m : 航线网络中的出发地数量;
- (3) N : 航线网络中的 $O-D$ 航程集合, $N = \{N^1, N^2, \dots, N^k\}$, N^k 表示以出发地 k 为起点的 $O-D$ 航程集合, $k = (1, 2, \dots, m)$;
- (4) n^k : N^k 中的 $O-D$ 航程数量;
- (5) C : 航线网络中所有航班的舱位集合, $C = \{C^1, C^2, \dots, C^k\}$, C^k 表示以出发地 k 为起点的全部 $O-D$ 航程中所有航班的舱位集合, $C^k = \{C^{1,k}, C^{2,k}, \dots, C^{jk}\}$, C^{jk} 表示以出发地 k 为起点的 $O-D$ 航程 j 中所有航班的舱位集合,

$j = (1, 2, \dots, n^k)$;

(6) c^{jk} : C^{jk} 中的舱位数量;

(7) t : 预订提前期中的预订观察点, $t = (0, 1, 2, \dots, T)$, 当 $t=0$ 时, 舱位预订开放, 当 $t=T$ 时, 航班起飞;

(8) Δ_t : 时刻 t 和 $t-1$ 之间的“预订提前期间隔”;

(9) I_u^{jk} : C^{jk} 中的舱位 i 在 Δ_t 中的历史预订开放状态, $I_u^{jk}=0$ 则表示舱位预订关闭, $I_u^{jk}=1$ 表示舱位预订开放, $i = (1, 2, \dots, c^{jk})$;

(10) S_t^{jk} : C^{jk} 在 Δ_t 中预订开放的舱位集合;

(11) $\bar{S}_t^{jk}$: C^{jk} 在 Δ_t 中预订关闭的舱位集合;

(12) B_u^{jk} : 系统记录到 S_t^{jk} 中的舱位 i 在 Δ_t 中的历史实名乘客成功预订需求量;

(13) D_u^{jk} : 系统记录到 $\bar{S}_t^{jk}$ 中的舱位 i 在 Δ_t 中的历史实名乘客未成功预订询问量;

(14) TS_u^{jk} : B_u^{jk} 和 D_u^{jk} 中记录到乘客购买或问询机票所属航程的中转联程状态, $TS_u^{jk}=0$ 表示直达航线, $TS_u^{jk}=1$ 表示中转联程航线;

(15) O_u^{jk} : B_u^{jk} 和 D_u^{jk} 中记录到乘客购买或问询机票所属航程的出发地, 当 $TS_u^{jk}=0$ 时, O_u^{jk} 所属航程出发地与 B_u^{jk} 和 D_u^{jk} 中记录到的保持一致;

(16) θ_u^{jk} : C^{jk} 中的舱位 i 在 Δ_t 中历史实名乘客潜在的取消预订比率;

(17) ρ_u^{jk} : C^{jk} 中的舱位 i 在 Δ_t 中未取消预订的历史实名乘客潜在的 no-shows 比率;

(18) P_u^{jk} : 在 Δ_t 中以 C^{jk} 中的舱位 i 为初次购买首选的历史实名乘客“初始需求”量;

(19) R_u^{jk} : 系统记录到由于初次购买所偏好的舱位预售关闭而选择转移并购买到 S_t^{jk} 中的舱位 i 的历史实名乘客“再现需求”量;

(20) RB_u^{jk} : 由于初次购买所偏好的舱位预售关闭而选择转移购买并最终预订成功的历史实名乘客, 他们对 $\bar{S}_t^{jk}$ 中的舱位 i 发出预订请求的“再现需求”重新预订量;

(21) SP_u^{jk} : 由于初次购买所偏好的舱位预售关闭而选择转移问询并最终选择放弃预订的历史实名乘客, 他们对 $\bar{S}_t^{jk}$ 中的舱位 i 发出预订请求的“溢出需求”重复询问量;

(22) CL_u^{jk} : C^{jk} 中的舱位 i 在 Δ_t 中历史实名乘客的取消预订量;

(23) NS_u^{jk} : C^{jk} 中的舱位 i 在 Δ_t 中未取消预订的历史实名乘客的 no-shows 量;

(24) U_u^{jk} : C^{jk} 中的舱位 i 在 Δ_t 中历史实名乘客实际“初始需求”无约束估计量。

其中, 符号(1)到(17)代表已知的历史实名需求信息。

2. 航线网络中基于乘客实名记录的“初始需求”分解查找法

为了从乘客实名记录中提取历史“初始需求”,在考虑网络替代效应的情况下,利用数据挖掘思想,本文所提分解查找法的具体步骤如下:

(1) 舱位预订开放状态检查

由 I_u^{jk} 、 B_u^{jk} 和 D_u^{jk} 的定义可知:当 $I_u^{jk}=1$ 时,预订系统记录到以出发地 k 为起点的 $O-D$ 航程 j 中舱位 i 在“预订提前期间隔” Δ_i 中的历史实名乘客成功预订需求量 B_u^{jk} ;当 $I_u^{jk}=0$ 时,表示实名需求受到存量控制策略限制而不能成功预订,预订系统仅能记录到历史实名乘客未成功预订的问询量 D_u^{jk} 。同时,根据 P_u^{jk} 、 R_u^{jk} 、 RB_u^{jk} 和 SP_u^{jk} 的定义可得到如下关系式:

$$B_u^{jk} = \begin{cases} P_u^{jk} + R_u^{jk}, & I_u^{jk} = 1 \\ 0, & I_u^{jk} = 0 \end{cases} \quad (1)$$

$$D_u^{jk} = \begin{cases} 0, & I_u^{jk} = 1 \\ P_u^{jk} + RB_u^{jk} + SP_u^{jk}, & I_u^{jk} = 0 \end{cases} \quad (2)$$

其中,当 $I_u^{jk}=1$ 时, P_u^{jk} 表示系统记录到 S_i^{jk} 中舱位 i 的成功预订“初始需求”量;当 $I_u^{jk}=0$ 时, P_u^{jk} 表示系统记录到 $\bar{S}_i^{jk}$ 中舱位 i 的未成功预订“初始需求”问询量;并且由式(1)和式(2)可得出如下结论:当 $I_u^{jk}=1$ 时, $P_u^{jk} = B_u^{jk} - R_u^{jk}$, $R_u^{jk} \geq 0$, $BR_u^{jk} = SP_u^{jk} = 0$;当 $I_u^{jk}=0$ 时, $P_u^{jk} = D_u^{jk} - RB_u^{jk} - SP_u^{jk}$, $R_u^{jk} = 0$, $BR_u^{jk} \geq 0$, $SP_u^{jk} \geq 0$ 。因此,在 $I_u^{jk}=1$ 和 $I_u^{jk}=0$ 的情况下,需要转到步骤(2)和(3),分别从所记录到的 B_u^{jk} 和 D_u^{jk} 中查找并确认出 R_u^{jk} 、 RB_u^{jk} 和 SP_u^{jk} 的大小,并最终确定 P_u^{jk} 的精确数值。

(2) 乘客实名需求信息的提取和机票中转联程状态检查

从各条 B_u^{jk} 和 D_u^{jk} 记录中分别提取乘客实名需求记录,不仅包括对乘客姓名、身份证号码以及预订和问询时刻 t' 等乘客信息的记录,还需确定所预订航班的出发地 k 、起飞时间 T 以及是否属于中转联程机票等航班信息。

(3) 乘客实名问询记录查询和 P_u^{jk} 的确定

A. 当 $I_u^{jk}=1$ 时

(a) 若某条 B_u^{jk} 记录中的 $TS_u^{jk}=0$,针对同一位实名乘客,查找其是否对 N^k 的所有 $O-D$ 航程在 $(T \pm 24)$ 小时内起飞的所有航班、从舱位开放预订到 t' 之间的所有预订关闭舱位存在历史问询记录。如果存在,则将该实名乘客在 B_u^{jk} 中的该条成功预订记录计入 R_u^{jk} ,否则,计入 P_u^{jk} 。

(b) 若某条 B_u^{jk} 记录中的 $TS_u^{jk}=1$,确定中转联程需求记录所属航程最初的出发点 $O_u^{jk}=K$,并判断其是否与同一位实名乘客在 B_u^{jk} 记录中所属出发点 k 相同,如果相同,分解查找法转到步骤(a)对 P_u^{jk} 进行计算;如果不同,则需针对同一位实名乘客、出发地 K 以及中转联程航段所属航程 J ,查找并判断该实名乘客是否对出发点 K 的所有 $O-D$ 航程 N^K 在 $(T \pm 24)$ 小时内起飞的所有航班、从舱位开放预订到

t' 之间所有预订关闭舱位存在历史问询记录,若存在,则将该实名乘客在 B_u^{jk} 中的成功预订记录计入 R_u^{JK} ,否则,计入 P_u^{JK} 。

B. 当 $I_u^{jk}=0$ 时

(a) 若某条 D_u^{jk} 记录中的 $TS_u^{jk}=0$,针对同一位实名乘客,查找其是否对 N^k 的所有 $O-D$ 航程中在 $(T\pm 24)$ 小时内起飞的所有航班、从 t' 到航班起飞时刻 T 的所有舱位存在成功预订或问询记录。如果存在,则进一步查找同一位实名乘客对 N^k 中在 $(T\pm 24)$ 小时内起飞的所有航班从舱位开放预订到 t' 之间的所有预订关闭舱位是否存在问询记录,如果同样存在,则将该实名乘客在 D_u^{jk} 中的问询记录计入 RB_u^{jk} ,否则,计入 P_u^{jk} 。同样地,如果判断出同一位实名乘客对 N^k 中在 $(T\pm 24)$ 小时内起飞的所有航班从 t' 到航班起飞时刻 T 的所有舱位不存在成功预订或问询记录,但该实名乘客对 N^k 中在 $(T\pm 24)$ 小时内起飞的所有航班从舱位开放预订到 t' 之间的所有预订关闭舱位存在问询记录,则将该实名乘客在 D_u^{jk} 中的问询记录计入 SP_u^{jk} ,否则,计入 P_u^{jk} 。

(b) 若某条 D_u^{jk} 记录中的 $TS_u^{jk}=1$,确定中转联程问询记录所属航程最初的出发点 $O_u^{jk}=K$,并判断其是否与同一位实名乘客在 D_u^{jk} 中所属出发点 k 相同,如果相同,分解查找法转到步骤(a)对 P_u^{jk} 进行计算;如果不同,则需针对同一位实名乘客、出发地 K 以及中转联程航段所属航程 J ,利用步骤(a)中的分解查找法查找并判断该实名乘客在出发点 K 的所有 $O-D$ 航程 N^K 中在 $(T\pm 24)$ 小时内起飞的所有航班“相关”舱位的预订和问询记录,并对该实名乘客在中转联程航程最初出发地 K 中的 P_u^{JK} 进行计算。

利用上述分解查找法,在 $I_u^{jk}=1$ 和 $I_u^{jk}=0$ 的情况下分别完成对各条乘客实名需求记录的查找和确认后,便得到了以出发地 k 为起点的 $O-D$ 航程 j 的舱位 i 在 Δ_i 中的 P_u^{jk} 。

3. 航线网络中历史实名乘客实际“初始需求”的无约束估计模型

在航空客运业中,航空公司经常会经历数量可观的乘客取消预订和 no-shows 情况,乘客可以在任何时间取消他们的预订,或者预订后根本就不出现。在大多数情况下,乘客不会遭受惩罚或罚款,在某些情况下,取消预订和 no-shows 的比例甚至高达 30% 以上^[12]。因此,在使用上节中所提分解查找法得到历史实名乘客“初始需求”量 P_u^{jk} 的具体数值后,需考虑乘客的取消预订和 no-shows 行为,最终完成对历史实名乘客实际“初始需求”的无约束估计计算。

(1) 当 $I_u^{jk}=1$ 时

在舱位预订开放状态下,由于预订系统能够跟踪 P_u^{jk} 中的实名乘客成功预订后的取消预订量 CL_u^{jk} 和 no-shows 量 NS_u^{jk} ,其实际“初始需求”应该从 P_u^{jk} 中将 CL_u^{jk} 和 NS_u^{jk} 剔除掉,其无约束估计模型如下所示:

$$U_u^{jk} = P_u^{jk} - CL_u^{jk} - NS_u^{jk} \quad (3)$$

(2) 当 $I_u^{jk} = 0$ 时

在舱位预订关闭状态下,因为乘客无法成功预订到相应舱位的机票,不会产生具体的取消预订和 no-shows 行为,但可以通过相同舱位和时间上以往成功预订机票的乘客的历史取消预订和 no-shows 情况对该情况下的实际“初始需求”进行计算。根据 θ_u^{jk} 、 ρ_u^{jk} 、 CL_u^{jk} 和 NS_u^{jk} 的定义可知: $CL_u^{jk} = \theta_u^{jk} P_u^{jk}$; $NS_u^{jk} = \rho_u^{jk} P_u^{jk} [1 - \theta_u^{jk}]$ 。

其中,在单航段和网络收益管理中,基于乘客历史实名记录分别对取消预订比率和 no-shows 比率进行计算的方法可参见 Morales 和 Wang^[10]、Garrow 和 Koppelman^[13]、Gorin^[14]、Neuling^[15]等的研究。因此,在确定 CL_u^{jk} 和 NS_u^{jk} 后,当舱位预订关闭时,历史实名乘客实际“初始需求”的无约束估计模型可表示如下:

$$\begin{aligned} U_u^{jk} &= P_u^{jk} - CL_u^{jk} - NS_u^{jk} \\ &= P_u^{jk} - \theta_u^{jk} P_u^{jk} - \rho_u^{jk} P_u^{jk} [1 - \theta_u^{jk}] \\ &= P_u^{jk} [1 - \theta_u^{jk}] [1 - \rho_u^{jk}] \end{aligned} \quad (4)$$

三、数值算例

假定国内某航空公司有如图 1 所示的航线网络,包含的城市(出发地)有北京(首都机场)、西安(咸阳机场)、青岛(流亭机场)和贵阳(龙洞堡机场)。该航线网络中以出发地 k 为起点的 $O-D$ 航程集合分别如表 1 所示。其中, $M = \{\text{北京、青岛、西安、贵阳}\}$, $m = 4$, $n^1 = 6$, $n^2 = 4$, $n^3 = 5$, $n^4 = 6$ 。由表 1 可知,该航空公司西安→青岛航程所考虑的航班舱位集合包括以下三种情景。

A. 以西安作为出发地($k=2$)的直达航程航线($j=2$)中所有航班的舱位集合 $C^{2,2}$;

B. 以西安作为出发地($k=2$)的中转联程航线($j=3$)中所有航段上(西安→北京、北京→青岛)所有航班的舱位集合 $C^{3,2}$;

C. 以贵阳作为出发地($k=4$)的中转联程航线($j=4$)中西安→青岛航段上所有航班的舱位集合 $C^{4,4}$ 。

根据图 1 和表 1 所示的航线网络,在西安→青岛某航班的第 4 个“预订提前期间隔” Δ_4 中编号 $i=3$ 的舱位,在不同的预订开放状态下,当其属于不同中转联程状态下的航班时,系统所记录到历史实名乘客成功预订需求量 $B_{3,4}^{i,k}$ 和未成功预订询问量 $D_{3,4}^{i,k}$ 如表 2 所示。利用本文所提的计算模型可以对在不同舱位预订开放状态以及不同机票中转联程状态下的历史实名乘客实际“初始需求” U_u^{jk} 进行无约束估计计算。

表 1 航线网络的 $O-D$ 航程集合

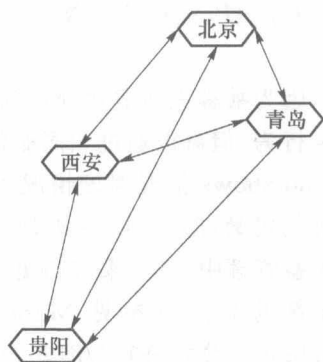


图 1 航线网络示意图

出发地 名称	出发地 编号 k	O-D 航程	
		集合 N^k	编号 j
北京	1	北京→青岛	1
		北京→青岛→贵阳	2
		北京→青岛→西安	3
		北京→西安	4
		北京→西安→贵阳	5
		北京→贵阳	6
西安	2	西安→北京	1
		西安→青岛	2
		西安→北京→青岛	3
		西安→贵阳	4
青岛	3	青岛→北京	1
		青岛→西安	2
		青岛→北京→西安	3
		青岛→贵阳	4
		青岛→西安→贵阳	5
贵阳	4	贵阳→北京	1
		贵阳→西安	2
		贵阳→西安→北京	3
		贵阳→西安→青岛	4
		贵阳→青岛	5
		贵阳→青岛→北京	6

首先,使用本文所提的分解查找法对各种情况下在 Δ_4 中编号 $i=3$ 的舱位历史实名乘客“再现需求” $R_{3,4}^{j,k}$ 、“再现需求”重新预订量 $RB_{3,4}^{j,k}$ 、“溢出需求”重复问询量 $SP_{3,4}^{j,k}$ 进行计算,并利用式(1)和式(2)得到“初始需求” $P_{3,4}^{j,k}$ 的具体数值。

然后,通过跟踪乘客的取消预订和 no-shows 记录,得到舱位预订开放状态下在 Δ_4 中编号 $i=3$ 的舱位历史实名乘客的取消预订量 $CL_{3,4}^{j,k}$ 和 no-shows 量 $NS_{3,4}^{j,k}$,并据此计算舱位预订关闭状态下历史实名乘客潜在的取消预订比例 $\theta_{3,4}^{j,k}$ 和 no-shows 比例 $\rho_{3,4}^{j,k}$,并得到该状态下历史实名乘客潜在的取消预订量 $CL_{3,4}^{j,k}$ 和 no-shows 量 $NS_{3,4}^{j,k}$ 。

最后,根据式(3)和式(4)分别计算各航程中实际“初始需求”的无约束估计值 $U_{3,4}^{j,k}$ 。

由表 2 可知,在考虑网络替代效应的情况下,本文所提基于乘客实名记录的分解查找法和无约束估计模型简单易行。

表 2 Δ_4 中编号 $i=3$ 的舱位在不同预订开放状态和航班中转联程状态下
历史实名乘客实际“初始需求”的无约束估计计算结果

$I_{3,4}^{j,k}$	出发地 k	航程 j	$TS_{3,4}^{j,k}$	$B_{3,4}^{j,k}$	$D_{3,4}^{j,k}$	分解查找法计算过程			$P_{3,4}^{j,k}$	无约束估计模型计算过程				$U_{3,4}^{j,k}$
						$R_{3,4}^{j,k}$	$RB_{3,4}^{j,k}$	$SP_{3,4}^{j,k}$		$\theta_{3,4}^{j,k}$	$CL_{3,4}^{j,k}$	$\rho_{3,4}^{j,k}$	$NS_{3,4}^{j,k}$	
1	2	2	0	15.00	0	5.00	0	0	10.00		2.00		1.00	7.00
	2	3	1	19.00	0	7.00	0	0	12.00		3.00		2.00	7.00
	4	4	1	28.00	0	10.00	0	0	18.00		6.00		3.00	9.00
0	2	2	0	0	35.00	0	12.00	11.00	12.00	0.20	2.40	0.13	1.25	8.35
	2	3	1	0	52.00	0	21.00	16.00	15.00	0.25	3.75	0.22	2.48	8.78
	4	4	1	0	63.00	0	29.00	20.00	14.00	0.33	4.62	0.25	2.35	7.04

四、结论

由于现有的无约束估计方法仅针对单航段收益管理实践,本文提出了一种考虑产品间网络替代效应的、使用乘客实名记录数据来判断和计算历史实际“初始需求”的分解查找法和无约束估计计算模型。所提方法为现有的数理统计类和选择模型类方法提供了一种可替换性的处理方式,并表现出如下新的特征,具有重要的现实意义:(1)综合考虑了网络替代效应下历史乘客对直达和中转联程航线产品的“初始需求”选择行为、“再现需求”重新预订行为、“溢出需求”重复问询行为以及乘客的取消预订和 no-shows 行为,并能根据所记录到的乘客实名数据对上述各类乘客需求进行查找、判断、分解和计算,能有效避免传统直接观察法对乘客“初始需求”的高估问题;(2)当公司从微观层面对网络环境下的乘客真实需求进行无约束

估计时,如果涉及的需求总量较小,数理统计类和选择模型类方法的准确性往往受到较大影响,本文所提方法的准确性不受数值大小的影响,在应用上显得更加灵活;(3) 为了在收益管理实践中对网络存量控制策略进行实时调整,本文所提方法消耗资源相对较少,适用于决策者基于大数据对未来“可实现需求(Realizable Demand)”(无约束需求预测值减去已实现需求量)的预测和更新。

但是,本文所提方法不完全适用于依靠价格手段来平衡供需的收益管理实践行业或公司,其成功应用需要将“价格遗憾”需求数据从乘客实名记录中分离出来,相较于系统内部所记录到由于存量控制策略而造成的“被拒绝”实名需求数据,在准确区分和判断乘客由于“价格遗憾”而决定放弃购买的选择行为方面,相关行业或公司可能存在一定的困难。同时,由于目前大多数收益管理系统无法对未事先预订而直接上门购买、并“被拒绝”或产生“价格遗憾”的 Walk-in 类随机实名需求进行准确的记录,本文所提方法并未提供针对上述需求类型的无约束估计方式;另外,在可获取乘客实名记录的情况下,有必要将本文所提网络无约束估计方法与传统单航段数理统计类和选择模型类方法在对公司收入影响等性能指标上进行仿真比较和评价。尽管存在上述局限性,但本文所提方法再次说明了乘客实名记录对于估计历史实际“初始需求”是非常有价值的,在考虑网络替代效应的同时,本研究的基本思想和计算逻辑能够被应用在更加广泛的网络收益管理实践中。

参考文献

- [1] Guo P, Xiao B C, Li J. Unconstraining methods in revenue management systems: Research overview and prospects [J]. Advances in Operations Research, 2012, 2012(270910).
- [2] Azadeh S S, Marcotte P, Savard G. A taxonomy of demand uncensoring methods in revenue management [J]. Journal of Revenue & Pricing Management, 2014, 13(6).
- [3] Vulcano G, van Ryzin G, Ratliff R. Estimating primary demand for substitutable products from sales transaction data [J]. Operations Research, 2012, 60(2).
- [4] van Ryzin G, Vulcano G. A market discovery algorithm to estimate a general class of nonparametric choice models [J]. Management Science, 2015, 61(2).
- [5] Ja S, Rao B V, Chandler S. Passenger recapture estimation in airline revenue management [C]// Proceedings of AGIFORS 41st Annual Symposium, Sydney: AGIFORS, 2001.
- [6] 郭鹏, 萧柏春, 李军. 收益管理中考虑顾客策略行为的多航班无约束估计方法[J]. 系统工程理论与实践, 2015, 35(5).