

不确定服务时间下分布式鲁棒门诊预约调度和排程

王珊珊, 李金林, 彭 春, 冉 伦

(北京理工大学管理与经济学院, 北京 100081)

摘要: 研究了单服务台门诊预约调度问题, 考虑了门诊服务中不确定服务时间, 基于服务时间联合概率分布的支撑集和矩等部分信息, 并利用平均绝对偏差刻画服务时间的相关性, 最小化最坏情况下期望等待成本和加班成本, 建立了分布式鲁棒优化门诊预约调度和排程模型, 利用优化方法结合对偶理论, 得出了模型的数学性质和易求解的等价问题, 确定患者最优的预约服务时长和服务次序. 最后, 仿真实验结果验证了模型的可行性和有效性.

关键词: 门诊预约调度; 排程; 不确定服务时间; 分布式鲁棒优化; 相关性

中图分类号: O224; F224

文献标识码: A

文章编号: 1000-5781(2019)04-0566-11

doi: 10.13383/j.cnki.jse.2019.04.011

Distributionally robust outpatient appointment scheduling and sequencing with uncertain service times

Wang Shanshan, Li Jinlin, Peng Chun, Ran Lun

(School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: This paper investigates outpatient appointment scheduling in a single server. Based on support and moment of random service time distribution, mean absolute deviation is employed to capture the correlation and distributionally robust outpatient appointment scheduling and sequencing models are proposed. Mathematical properties and tractable counterparts of the models are derived by optimization methods and duality theory, so as to determine the service sequence and time allowance for each appointment. Finally, numerical experiments are performed and the results indicate the feasibility and effectiveness of the model.

Key words: outpatient appointment scheduling; sequencing; uncertain service time; distributionally robust optimization; correlation

1 引言

门诊是医疗服务体系中的关键资源之一, 也是医院服务的一个重要的窗口, 具有门诊患者数量巨大、资源供不应求等特点^[1]. 作为“看病难”最突出的部门, 其服务质量直接决定着医院的服务水平, 门诊资源的利用效率直接影响着后续部门的运作效率. 因此, 国家出台了一系列的政策, 旨在开展并推进预约诊疗服务, 有针对性地解决群众反映突出的“看病难”问题. 随着一系列政策的出台和互联网技术的迅速发展, 门诊预约服务在全国各地得到了广泛的推行并逐渐完善, 预约挂号服务已经成为医院提高门诊资源利用效率、解决门诊服务难题的重要切入点, 门诊预约调度也成为医疗运作管理的一个研究热点.

收稿日期: 2016-09-01; 修订日期: 2017-05-25.

基金项目: 国家自然科学基金重点资助项目(71432002); 国家自然科学基金资助项目(71672011).

门诊预约调度主要包括两个阶段, 第一阶段为排程(sequencing), 给定不同类型患者的随机服务时间, 确定最优的预约服务次序. 第二阶段为调度(scheduling), 即在门诊排程固定的前提下, 确定患者最优的预约服务时长. 科学合理的门诊预约调度和排程策略能有效的减少患者的等待时间和医生的加班时间, 提高门诊服务的质量和患者的满意度^[2]. 一个非常关键的问题是门诊预约调度过程中服务时间是不确定的, 服务时间的不确定性不仅增加了门诊预约调度难度, 也在一定程度上影响了门诊的服务质量和资源的利用效率. 因此, 如何有效的度量门诊预约调度过程中服务时间的不确定性, 成为学者们研究的关键.

随机规划和分布式鲁棒优化是解决不确定问题常用的方法. 现阶段基于随机规划的门诊预约调度的研究较多. Begen和Queyranne^[3]假设服务时间的联合离散概率分布已知且服务时间是相关的, 得出了最优的门诊调度策略. Ge等^[4]在文献[3]的基础上, 将其结果推广到分段线性成本函数. Castaing等^[5]基于同时服务多个患者的情况, 建立了一个两阶段随机整数规划模型. 姜博文等^[6]研究了门诊预约能力分配问题, 建立了一个门诊净收益的期望和方差模型. Denton等^[7]指出当存在两种不同类型的患者且服务时间独立时, 按照服务时间方差递增(order of variance, OV)的服务次序最优. Mak等^[8]基于预约排程模型与随机库存模型结构的相关性, 设计了一种新的启发式算法, 得到了预约排程的近似最优解.

随机规划的门诊预约调度一般依赖于服务时间的概率分布已知, 然而患者的就诊数据往往不能够准确地描述服务时间的概率分布. 现有的研究为了降低模型求解的复杂性, 大多假设服务时间是相互独立的, 当医生连续服务多个患者时, 会导致服务时间相互影响, 忽略服务时间相关性的门诊预约调度和排程, 应用在医院实际运作中可能产生较大偏差. 预约系统的排程一直是研究的难点, 部分学者只能采用启发式算法得到最优预约次序的近似解. 因此, 本文采用分布式鲁棒优化的方法研究门诊预约调度和排程问题.

分布式鲁棒优化方法是一种有效处理随机变量分布信息不完全的决策方法, 考虑最坏情况下的策略, 给出具有鲁棒性的决策, 同时可以利用随机变量的部分分布信息刻画随机变量的相关性. 作为一种有效处理不确定性的工具, 分布式鲁棒优化方法得到了越来越多学者的关注, 并在运营管理领域得到了广泛的运用^[9-11], 但现阶段关于分布式鲁棒优化的预约调度和排程的研究较少. Kong等^[12]研究了单服务台的预约调度问题, 基于服务时间概率分布的均值和协方差矩阵, 利用协方差矩阵体现服务时间的相关性, 建立了分布式鲁棒优化预约调度模型并得到了近似最优解. Mak等^[13]在文献[12]的基础上, 基于服务时间概率分布一阶矩和二阶矩信息, 建立了分布式鲁棒预约调度和排程模型, 在一定条件下, 给出了服务时长的闭式最优解并证明服务时间方差递增的服务次序最优. 考虑服务时间概率分布的支撑集和一阶矩信息, 得到了相似的结论. Qi^[14]考虑服务时间不确定, 以缓解不公平性和患者由于等待引起的不满为优化目标, 分别建立随机模型和分布式鲁棒模型, 在分布式鲁棒模型中构建了统计量刻画服务时间的相关性.

本文在Mak等^[13]工作的基础上, 进一步地, 同时考虑服务时间概率分布的支撑集、一阶矩、平均绝对偏差及服务时间的相关性, 并分析了不同的相关性系数对最优的门诊预约调度和排程策略的影响. 文章主要完成了以下几个部分工作: 1) 考虑服务时间的不确定性, 基于服务时间的支撑集和矩等部分信息, 并利用绝对平均偏差刻画患者服务时间的相关性, 分别建立分布式鲁棒优化门诊预约调度和排程模型. 2) 由于所建分布式模型为非线性的双层 min-max 或 max-min 问题, 充分结合分布式鲁棒模型的数学性质, 将其转化为易求解的等价问题, 最后确定患者最优的预约服务时长和服务次序. 3) 在实验设计部分, 分别得到了同种类型患者的最优调度策略和不同种类型患者的最优调度和排程策略. 本文构建的分布式鲁棒优化模型, 在一定程度上丰富了基于分布式鲁棒优化理论的门诊预约调度文献, 也为医院管理者提供相关的决策支持.

2 分布式鲁棒优化预约调度和排程模型

2.1 问题描述

本文考虑单服务台问题, 假设门诊同一时间只有一个医生提供服务, 一个医生一次只能服务一个患者; 每天安排的患者人数固定, 患者准时到达, 不存在迟到、爽约等行为; 医生不存在迟到、临时停诊等行为; 服

务时间精确的联合概率分布未知, 仅已知服务时间的支撑集、一阶矩及平均绝对偏差的上界. 基于此, 研究不同患者类型和服务时间不确定下的门诊预约调度问题, 确定患者最优的预约服务时长和服务次序, 建立两个分布式鲁棒优化的模型. 2.2 节假设预约服务次序固定, 建立一个分布式鲁棒优化的预约调度模型研究最优的预约服务时长. 2.3 节建立分布式鲁棒优化预约排程模型研究最优的预约服务时长和服务次序.

符号说明具体为:

参数与集合: $N = \{1, 2, \dots, n\}$: 一个周期内预约患者的集合, $k, i \in N$; $\tilde{s}_k$ 为非负随机变量, 表示第 k 个患者不确定服务时间; μ_k 表示第 k 个患者不确定服务时间的期望; σ_k 表示第 k 个患者不确定服务时间平均绝对偏差的上界; $D_{\tilde{s}_k}$ 表示第 k 个患者不确定服务时间的支撑集, $D_{\tilde{s}_k} = [\underline{s}_k, \bar{s}_k]$; P 表示满足分布约束条件的随机向量 $\tilde{s} = (\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_n)$ 的联合概率分布; Γ 为联合概率分布 P 的集合; ω_k 为第 k 个患者的等待时间; o 为医生的加班时间; c_k 为第 k 个患者的等待成本; γ 为医生的加班成本; T 为一个服务周期内预约时间长度. 决策变量: x_k 为第 k 个患者的预约服务时长; y_{ik} 为 0-1 分配变量, 如果第 i 个患者被分配到第 k 个位置, 则 $y_{ik} = 1$, 否则为 0.

2.2 分布式鲁棒优化预约调度模型

患者按照序列 $1, 2, \dots, n$ 的顺序就诊, 给定患者的预约服务时长 $\mathbf{x}$ 和随机服务时间 $\tilde{s}$, 则第 1 个患者的等待时间 $\omega_1 = 0$, 第 $k+1$ 个患者的等待时间和医生的加班时间为

$$\omega_{k+1} = \max\{\omega_k + \tilde{s}_k - x_k, 0\}, \quad k = 1, 2, \dots, n-1 \quad (1)$$

$$o = \max\{\omega_n + \tilde{s}_n - x_n, 0\}, \quad (2)$$

患者的等待成本和医生的加班成本为 $f(\tilde{s}, \mathbf{x}) = \sum_{k=2}^n c_k \omega_k + \gamma o$, 由于式 (1) 和式 (2) 含有最大化问题, $f(\tilde{s}, \mathbf{x})$ 不易求解. 因此将 ω, o 松弛, 引入 ω', o' 表示松弛后变量, 可得 $f(\tilde{s}, \mathbf{x})$ 另一种表达式, 即为问题式 (3)~式 (7) 的最优目标值, 即

$$f(\tilde{s}, \mathbf{x}) = \text{Min}_{\omega', o'} \sum_{k=2}^n c_k \omega'_k + \gamma o' \quad (3)$$

$$\text{s.t. } \omega'_2 \geq \tilde{s}_1 - x_1 \quad (4)$$

$$\omega'_{k+1} \geq \omega'_k + \tilde{s}_k - x_k, \quad k = 2, 3, \dots, n-1 \quad (5)$$

$$o' \geq \omega'_n + \tilde{s}_n - x_n \quad (6)$$

$$\omega'_k, o' \geq 0, \quad k = 2, 3, \dots, n. \quad (7)$$

根据对偶理论, 将问题式 (3)~式 (7) 等价于最大化问题式 (8)~式 (11), 即

$$f(\tilde{s}, \mathbf{x}) = \text{Max}_l \sum_{k=1}^n (\tilde{s}_k - x_k) l_k \quad (8)$$

$$\text{s.t. } l_k - l_{k-1} \geq -c_k, \quad k = 2, 3, \dots, n \quad (9)$$

$$l_k \leq \gamma, \quad k = n \quad (10)$$

$$l_k \geq 0, \quad k = 1, 2, \dots, n, \quad (11)$$

其中 $\mathbf{l} = (l_1, l_2, \dots, l_n)$ 为对偶变量.

已知服务时间的矩信息、支撑集和平均绝对偏差的上界. 由于在单服务台问题中, 服务时间相互影响, 具体为医生在提供服务的过程中, 当服务时间过长, 即偏离平均服务时间较多时, 医生往往加快诊疗速度, 以实现预期安排的工作量和总服务时间, 导致接下来的服务时间小于平均服务时间情况的发生, 当服务时间过短时则情况相反, 同时这种偏离程度的判断与随机服务时间的波动有关, 本文不考虑单个服务时间之间的相关性, 所以用 $E_P\left(\left|\sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k}\right|\right) \leq \delta$ 表示服务时间相关性的约束^[14],

$0 \leq \delta \leq n$, δ 越小相关性越强, δ 越大相关性越弱. 服务时间的联合概率分布 P 的不确定集合: $\Gamma = \{P | E_P(\tilde{s}_k) = \mu_k, P(\tilde{s}_k \in [\underline{s}_k, \bar{s}_k]) = 1, E_P(|\tilde{s}_k - \mu_k|) \leq \sigma_k, E_P\left(\left|\sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k}\right|\right) \leq \delta\}$, 可得基于期望测度的预约调度分布式鲁棒优化模型

$$\min_{\mathbf{x} \in \Omega} \max_{P \in \Gamma} E_P(f(\tilde{\mathbf{s}}, \mathbf{x})), \quad (12)$$

其中 $\Omega = \{\mathbf{x} \geq 0, \sum_{k=1}^n x_k = T\}$.

预约调度模型 (12) 可看作一个两阶段优化问题: 首先给定一个决策变量 $\mathbf{x}$ 的值, 求得期望总成本的最大值. 然后在决策变量 $\mathbf{x}$ 的约束下, 求得期望总成本所有最大值中的最小值, 从而得到分布式鲁棒优化预约调度问题的最优解. 首先分析模型 (12) 的内层最大化问题, 即给定一个决策变量 $\mathbf{x}$ 的值, 求得期望总成本的最大值. 内层最大化问题

$$\max_{P \in \Gamma} E_P(f(\tilde{\mathbf{s}}, \mathbf{x})). \quad (13)$$

引理 1 对于任意给定的 $\mathbf{x}$, 内层优化问题 (13) 等价于问题式 (14)~式 (16), 即

$$\min \left(\rho_0 + \sum_{k=1}^n \mu_k \alpha_k + \sum_{k=1}^n \sigma_k \beta_k + \delta \phi \right) \quad (14)$$

$$\text{s.t. } \rho_0 \geq \max_{l \in \Lambda} h(l, \alpha, \beta, \phi) \quad (15)$$

$$\beta_k, \phi \geq 0, \quad k = 1, 2, \dots, n, \quad (16)$$

其中 $h(l, \alpha, \beta, \phi) = \max_{\tilde{\mathbf{s}} \in D_{\tilde{\mathbf{s}}}} \left[\sum_{k=1}^n [(\tilde{s}_k - x_k)l_k - \alpha_k \tilde{s}_k - \beta_k |\tilde{s}_k - \mu_k|] - \phi \left| \sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right]$, Λ 为式 (9)~式 (11) 定义的关于 l 的可行域.

证明 内层最大化问题 (13) 的对偶问题为问题式 (17)~式 (19), 即

$$\min \left(\rho_0 + \sum_{k=1}^n \mu_k \alpha_k + \sum_{k=1}^n \sigma_k \beta_k + \delta \phi \right) \quad (17)$$

$$\text{s.t. } \rho_0 \geq f(\tilde{\mathbf{s}}, \mathbf{x}) - \sum_{k=1}^n \alpha_k \tilde{s}_k - \sum_{k=1}^n \beta_k |\tilde{s}_k - \mu_k| - \phi \left| \sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right|, \quad \forall \tilde{s}_k \in [\underline{s}_k, \bar{s}_k], \quad k = 1, 2, \dots, n \quad (18)$$

$$\beta_k, \phi \geq 0, \quad k = 1, 2, \dots, n, \quad (19)$$

其中 $\rho_0, \alpha_k, \beta_k, \phi$ 是原问题 (13) 的对偶变量.

由强对偶理论知原问题和对偶问题等价. 由于约束式 (16) 中对所有的 $\tilde{s}_k \in [\underline{s}_k, \bar{s}_k]$, $k = 1, 2, \dots, n$ 都有不等式成立, 即考虑约束 $\rho_0 \geq \max_{\tilde{\mathbf{s}} \in D_{\tilde{\mathbf{s}}}} \left[f(\tilde{\mathbf{s}}, \mathbf{x}) - \sum_{k=1}^n \alpha_k \tilde{s}_k - \sum_{k=1}^n \beta_k |\tilde{s}_k - \mu_k| - \phi \left| \sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right]$.

进一步有

$$\begin{aligned} & \max_{\tilde{\mathbf{s}} \in D_{\tilde{\mathbf{s}}}} \left[f(\tilde{\mathbf{s}}, \mathbf{x}) - \sum_{k=1}^n \alpha_k \tilde{s}_k - \sum_{k=1}^n \beta_k |\tilde{s}_k - \mu_k| - \phi \left| \sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right] \\ &= \max_{\tilde{\mathbf{s}} \in D_{\tilde{\mathbf{s}}}} \left[\max_{l \in \Lambda} \sum_{k=1}^n (\tilde{s}_k - x_k) l_k - \sum_{k=1}^n \alpha_k \tilde{s}_k - \sum_{k=1}^n \beta_k |\tilde{s}_k - \mu_k| - \phi \left| \sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right] \\ &= \max_{\tilde{\mathbf{s}} \in D_{\tilde{\mathbf{s}}}} \max_{l \in \Lambda} \left[\sum_{k=1}^n [(\tilde{s}_k - x_k) l_k - \alpha_k \tilde{s}_k - \beta_k |\tilde{s}_k - \mu_k|] - \phi \left| \sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right] \\ &= \max_{l \in \Lambda} \max_{\tilde{\mathbf{s}} \in D_{\tilde{\mathbf{s}}}} \left[\sum_{k=1}^n [(\tilde{s}_k - x_k) l_k - \alpha_k \tilde{s}_k - \beta_k |\tilde{s}_k - \mu_k|] - \phi \left| \sum_{k=1}^n \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right], \end{aligned} \quad (20)$$

所以引理 1 成立.

由于约束式 (15) 中含有 Max 问题, 仍然不易求解, 因此考虑如下最大化问题

$$\text{Max}_{l \in \Lambda} h(l, \alpha, \beta, \phi). \quad (21)$$

对于固定的 α, β 和 ϕ , 易证 $h(l, \alpha, \beta, \phi)$ 关于 l 是凸函数, 即 (21) 是一个最大化凸函数问题, 求解比较复杂. 在处理这个最大化问题时, 主要借鉴文献 [13] 处理类似问题的思想.

定义变量 $\pi_{kj} = \sum_{i=k+1}^j c_i$, 当 $1 \leq k \leq j \leq n$ 时, $\pi_{kj} = \gamma + \sum_{i=k+1}^n c_i$, 当 $1 \leq k \leq n, j = n+1$ 时, 其中定义 $\sum_{i=n+1}^n c_i = 0$. 由于 (19) 在可行域 Λ 的极值点处取得最优值, 对于区间 $[w, j] \subseteq [1, n+1]$, 如果任意的 $k \in [w, j]$, $l_k = 0$ 当且仅当 $k = j$, 则有 $l_k = \pi_{kj}$. 引入 0-1 变量 t_{wj} , 如果区间 $[w, j]$ 满足上述性质则 $t_{wj} = 1$, 否则为 0. 此时有 $\sum_{w=1}^i \sum_{j=i}^{n+1} t_{wj} = 1, i = 1, 2, \dots, n+1$, 当 $t_{wj} = 1$ 时,

$$\begin{aligned} & \sum_{k=w}^j [(\tilde{s}_k - x_k)l_k - \alpha_k \tilde{s}_k - \beta_k |\tilde{s}_k - \mu_k|] - \phi \left| \sum_{k=w}^j \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \\ &= \sum_{k=w}^j [(\tilde{s}_k - x_k)\pi_{kj} - \alpha_k \tilde{s}_k - \beta_k |\tilde{s}_k - \mu_k|] - \phi \left| \sum_{k=w}^j \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right|. \end{aligned} \quad (22)$$

令 $\varphi_{w,j}(\pi_{kj}, \alpha_k, \beta_k, \phi) = \text{Max}_{\tilde{s} \in D_{\tilde{s}}} \left[\sum_{k=w}^j [(\tilde{s}_k - x_k)\pi_{kj} - \alpha_k \tilde{s}_k - \beta_k |\tilde{s}_k - \mu_k|] - \phi \left| \sum_{k=w}^j \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right], 1 \leq w \leq n+1, w \leq j \leq n+1$, 所以问题 (21) 等价于问题式 (23)~式 (25), 即

$$\text{Max}_t \sum_{w=1}^{n+1} \sum_{j=w}^{n+1} \varphi_{w,j}(\pi_{kj}, \alpha_k, \beta_k, \phi) t_{wj} \quad (23)$$

$$\text{s.t.} \sum_{w=1}^i \sum_{j=i}^{n+1} t_{wj} = 1, \quad i = 1, 2, \dots, n+1 \quad (24)$$

$$t_{wj} \in \{0, 1\}, \quad 1 \leq w \leq j \leq n+1. \quad (25)$$

问题式 (23)~式 (25) 是最大化 t 的线性规划问题, 由于约束式 (24) 是完全幺模的, 则可将 0-1 变量 t_{wj} 松弛为 $t_{wj} \geq 0$, 最优解不变. 根据对偶理论问题式 (23)~式 (25) 等价于问题式 (26)~式 (27), 即

$$\text{Min}_{\lambda} \sum_{k=1}^n \lambda_k \quad (26)$$

$$\text{s.t.} \sum_{k=w}^{\min\{j,n\}} \lambda_k \geq \varphi_{w,\min\{j,n\}}(\pi_{kj}, \alpha_k, \beta_k, \phi), \quad 1 \leq w \leq n, w \leq j \leq n+1, \quad (27)$$

所以对于任意给定的 $x \in \Omega$ 内层优化问题 (13) 等价于问题式 (28)~式 (31), 即

$$\text{Min} \left(\rho_0 + \sum_{k=1}^n \mu_k \alpha_k + \sum_{k=1}^n \sigma_k \beta_k + \delta \phi \right) \quad (28)$$

$$\text{s.t.} \rho_0 \geq \sum_{k=1}^n \lambda_k \quad (29)$$

$$\sum_{k=w}^{\min\{j,n\}} \lambda_k \geq \varphi_{w,\min\{j,n\}}(\pi_{kj}, \alpha_k, \beta_k, \phi), \quad 1 \leq w \leq n, w \leq j \leq n+1 \quad (30)$$

$$\beta_k, \phi \geq 0, \quad k = 1, 2, \dots, n. \quad (31)$$

定理 1 预约调度模型 (12) 与问题式 (32)~式 (40) 等价, 即

$$\text{Min} \left(\rho_0 + \sum_{k=1}^n \mu_k \alpha_k + \sum_{k=1}^n \sigma_k \beta_k + \delta \phi \right) \quad (32)$$

$$\text{s.t. } \rho_0 \geq \sum_{k=1}^n \lambda_k \quad (33)$$

$$\sum_{k=w}^{\min\{j,n\}} \lambda_k \geq \sum_{k=w}^{\min\{j,n\}} \xi_k^{w,j}, \quad 1 \leq w \leq n, w \leq j \leq n+1 \quad (34)$$

$$\xi_k^{w,j} \geq (\bar{s}_k u_k^{w,j} - \underline{s}_k v_k^{w,j}) + \mu_k (z_k^{w,j} - q_k^{w,j}) + \mu_k (a^{w,j} - b^{w,j}) / \sigma_k - x_k \pi_{kj}, \quad (35)$$

$$1 \leq w \leq n, w \leq j \leq n+1, w \leq k \leq \min\{j, n\}$$

$$u_k^{w,j} - v_k^{w,j} + z_k^{w,j} - q_k^{w,j} + (a^{w,j} - b^{w,j}) / \sigma_k \geq \pi_{kj} - \alpha_k, \quad (36)$$

$$1 \leq w \leq n, w \leq j \leq n+1, w \leq k \leq \min\{j, n\}$$

$$z_k^{w,j} + q_k^{w,j} \leq \beta_k, \quad 1 \leq w \leq n, w \leq j \leq n+1, w \leq k \leq \min\{j, n\} \quad (37)$$

$$a^{w,j} + b^{w,j} \leq \phi, \quad 1 \leq w \leq n, w \leq j \leq n+1 \quad (38)$$

$$\sum_{k=1}^n x_k = T \quad (39)$$

$$u_k^{w,j}, v_k^{w,j}, z_k^{w,j}, q_k^{w,j}, a^{w,j}, b^{w,j}, \beta_k, \phi, x_k \geq 0, \quad 1 \leq w \leq n, w \leq j \leq n+1, w \leq k \leq n. \quad (40)$$

证明 在约束式 (30) 中

$$\varphi_{w, \min\{j, n\}}(\pi_{kj}, \alpha_k, \beta_k, \phi) = \max_{\tilde{s} \in D_{\tilde{s}}} \left[\sum_{k=w}^{\min\{j, n\}} [(\tilde{s}_k - x_k) \pi_{kj} - \alpha_k \tilde{s}_k - \beta_k |\tilde{s}_k - \mu_k|] - \phi \left| \sum_{k=w}^{\min\{j, n\}} \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right| \right],$$

令 $\theta_k = |\tilde{s}_k - \mu_k|$, $\eta = \left| \sum_{k=w}^j \frac{\tilde{s}_k - \mu_k}{\sigma_k} \right|$, 根据对偶理论, 式 (28)~式 (31) 等价于问题式 (41)~式 (45), 即

$$\text{Min} \sum_{k=w}^{\min\{j, n\}} [(\bar{s}_k u_k - \underline{s}_k v_k) + \mu_k (z_k - q_k) + \mu_k (a - b) / \sigma_k - x_k \pi_{kj}] \quad (41)$$

$$\text{s.t. } u_k - v_k + z_k - q_k + (a - b) / \sigma_k \geq \pi_{kj} - \alpha_k, \quad w \leq k \leq \min\{j, n\} \quad (42)$$

$$z_k + q_k \leq \beta_k, \quad w \leq k \leq \min\{j, n\} \quad (43)$$

$$a + b \leq \phi \quad (44)$$

$$u_k, v_k, z_k, q_k, a, b \geq 0, \quad w \leq k \leq \min\{j, n\}, \quad (45)$$

又因为 $\varphi_{w, \min\{j, n\}}(\pi_{kj}, \alpha_k, \beta_k, \phi)$ 的值与 w 和 j 有关, 引入中间变量 $\xi_k^{w,j}$, 则有定理 1 成立.

当预约服务次序固定时, 根据定理 1 分布式鲁棒优化预约调度模型 (12) 可转化成易求解的线性规划问题, 进而可以确定患者的最优的预约服务时长, 最小化最坏分布情况下的期望总成本.

2.3 分布式鲁棒优化预约排程模型

对于不同就诊类型的患者 i 和不同的门诊服务位置 k , 决策变量 y_{ik} 满足约束 $\sum_{i=1}^n y_{ik} = 1$, $\sum_{k=1}^n y_{ik} = 1$, 分别表示每个位置只能安排一个患者, 每个患者只能被安排在一个位置. 则第 k 个位置随机服务时间为 $s_k(y) = \sum_{i=1}^n \tilde{s}_i y_{ik}$, 给定每个位置的预约时间长度 x_k 和随机服务时间 $s_k(y)$, 总的等待成本和加班成本 $f_1(\tilde{s}, x)$ 为问题式 (46)~式 (50) 的最优目标值, 即

$$f_1(\tilde{s}, \mathbf{x}) = \text{Min}_{\omega', o'} \sum_{k=2}^n c_k \omega'_k + \gamma o' \quad (46)$$

$$\text{s.t. } \omega'_2 \geq s_1(y) - x_1 \quad (47)$$

$$\omega'_{k+1} \geq \omega'_k + s_k(y) - x_k, \quad k = 2, 3, \dots, n-1 \quad (48)$$

$$o' \geq \omega'_n + s_n(y) - x_n \quad (49)$$

$$\omega'_k, o' \geq 0, \quad k = 2, 3, \dots, n. \quad (50)$$

分布式鲁棒优化预约排程模型为

$$\text{Min}_{\mathbf{x} \in \Omega, \mathbf{y} \in \Upsilon} \text{Max}_{P \in \Gamma} E_P(f_1(\tilde{s}, \mathbf{x})), \quad (51)$$

其中 $\Omega = \{\mathbf{x} \geq 0, \sum_{k=1}^n x_k = 1\}$, $\Upsilon = \{\mathbf{y} \in \{0, 1\}, \sum_{i=1}^n y_{ik} = 1, \sum_{k=1}^n y_{ik} = 1\}$, $\Gamma = \{P | E_P(\tilde{s}_i) = \mu_i, P(\tilde{s}_i \in [\underline{s}_i, \bar{s}_i]) = 1, E_P(|\tilde{s}_i - \mu_i|) \leq \sigma_i, E_P\left(\left|\sum_{i=1}^n \frac{\tilde{s}_i - \mu_i}{\sigma_i}\right|\right) \leq \delta\}$.

同 2.2 中类似的证明过程, 模型 (51) 的最终等价形式为问题式 (52)~式 (60), 即

$$\text{Min} \left(\rho_0 + \sum_{i=1}^n \mu_i \alpha_i + \sum_{i=1}^n \sigma_i \beta_i + \delta \phi \right) \quad (52)$$

$$\text{s.t. } \rho_0 \geq \sum_{k=1}^n \lambda_k \quad (53)$$

$$\sum_{k=w}^{\min\{j, n\}} \lambda_k \geq \xi^{w, j}, \quad 1 \leq w \leq n, \quad w \leq j \leq n+1 \quad (54)$$

$$\begin{aligned} \xi^{w, j} \geq & \left(\sum_{i=1}^n \bar{s}_i u_i^{w, j} - \sum_{i=1}^n \underline{s}_i v_i^{w, j} \right) + \sum_{i=1}^n \mu_i (z_i^{w, j} - q_i^{w, j}) + \\ & \sum_{i=1}^n \mu_i (a^{w, j} - b^{w, j}) / \sigma_i - \sum_{k=w}^{\min\{j, n\}} x_k \pi_{kj}, \quad 1 \leq w \leq n, \quad w \leq j \leq n+1 \end{aligned} \quad (55)$$

$$u_i^{w, j} - v_i^{w, j} + z_i^{w, j} - q_i^{w, j} + (a^{w, j} - b^{w, j}) / \sigma_i \geq \sum_{k=w}^{\min\{j, n\}} (y_{ik} \pi_{kj} - \frac{1}{n} \alpha_i), \quad (56)$$

$$1 \leq i \leq n, \quad 1 \leq w \leq n, \quad w \leq j \leq n+1$$

$$z_i^{w, j} + q_i^{w, j} \leq \sum_{k=w}^{\min\{j, n\}} \beta_i / n, \quad 1 \leq i \leq n, \quad 1 \leq w \leq n, \quad w \leq j \leq n+1 \quad (57)$$

$$a^{w, j} + b^{w, j} \leq \sum_{k=w}^{\min\{j, n\}} \phi / n, \quad 1 \leq w \leq n, \quad w \leq j \leq n+1 \quad (58)$$

$$\sum_{k=1}^n x_k = T, \quad \sum_{i=1}^n y_{ik} = 1, \quad \sum_{k=1}^n y_{ik} = 1 \quad (59)$$

$$u_i^{w, j}, v_i^{w, j}, z_i^{w, j}, q_i^{w, j}, a^{w, j}, b^{w, j}, \beta_k, \phi, x_k \geq 0, y_{ik} \in \{0, 1\}, \quad 1 \leq i \leq n, \quad 1 \leq w \leq n, \quad w \leq j \leq n+1. \quad (60)$$

3 算例分析

算例分析分为两部分, 首先, 预约服务次序固定, 研究同种就诊类型患者最优的门诊预约服务时长. 其次, 存在不同就诊类型患者时, 研究最优的门诊预约服务时长和服务次序. 部分参数设置为: 服务时间的支

撑集 $D_s = [5, 30]$, 一个服务周期内预约时间长度 $T = \sum_{k=1}^n \mu_k + R \sqrt{\sum_{k=1}^n \sigma_k^2}$, 其中 R 是调节预约时间长度的参数^[13], 病人的等待成本 $c_k = 1, \forall k = 2, 3, \dots, n$, 所以医生的加班成本 γ 即为医生与患者的相对时间成本.

3.1 同种类型患者的门诊预约调度

在预约服务次序固定的情况下, 研究同种类型患者的门诊预约调度策略. 假设同种类型患者的服务时间服从相同概率分布. 具体参数设置: 患者的平均服务时间 $\mu = 15 \text{ min}$, 平均绝对偏差的上界 $\sigma = 3$, 本节中, 变化一个周期内预约患者的人数 $n = 7 : 1 : 10$, $\delta = 3 : 1 : 6$, 医生与患者的相对时间成本 $\gamma = 2 : 1 : 5$, $R = -0.5, 0, 0.5$. 求得不同患者数量下的最优预约服务时长以及不同 δ 下的最优预约服务时长分别如图1和图2所示.

当预约服务次序固定时, 基于分布式鲁棒优化预约调度模型求解的最优预约服务时长呈现出递减的趋势, 如图1和图2所示. 主要的原因是当序列前面的患者就诊出现延迟现象时, 会对该患者之后的所有患者的实际就诊时间产生影响, 即序列前面的延迟影响较大. 考虑最坏分布情况下的最优服务时长时, 为了减少由于序列前面延迟产生的影响, 在实际安排就诊时长时, 序列前面的预约服务时长应较长, 序列后面的预约服务时长应较短, 给患者安排的预约服务时长随预约次序呈现出递减的趋势. 同时由于服务时间相关性的约束, 在最坏分布情况下, 序列开始的随机服务时间较长, 此时医生会增加诊疗速度, 从而最优预约服务时长出现波动. 图1中存在不同患者数量时, 患者数量越少, 序列前面患者的服务时间越长, 序列后面患者的服务时间越短. 当 δ 相同时, 患者数量越少, 服务时间概率分布的不确定集合范围越大, 预约调度越保守, 从而序列前面患者的预约服务时间越长, 序列后面患者的预约服务时间越短. 同理考虑对 δ 的不同约束, 当 δ 越大时, 服务时间总体的相关性越弱, 此时分布的不确定集合范围也越大, 从而序列前面分配的时间越长, 波动越大. 具体见图2.

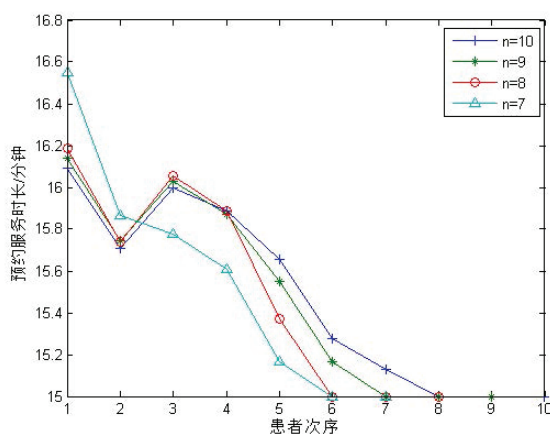


图1 不同患者数量下最优的预约服务时长

Fig. 1 The optimal appointment time allowances with different number of patients

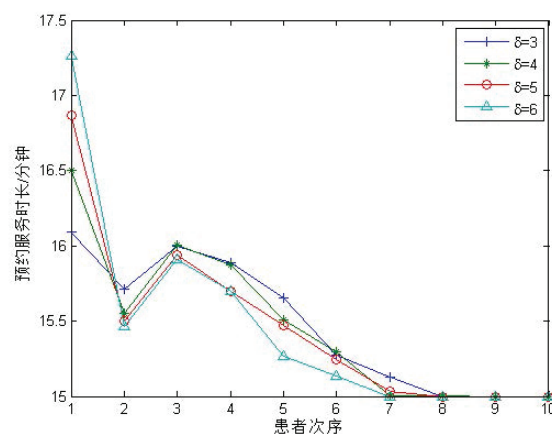


图2 不同 δ 下最优的预约服务时长

Fig. 2 The optimal appointment time allowances with different

医生的超时成本通常会高于患者的等待成本, 由于不同的门诊类型, 如普通门诊和专家门诊的超时成本通常不同, 所以图3给出了医生与患者不同相对时间成本下的最优预约服务时长. 由于序列后面对加班时间的影响较大, 所以当相对时间成本较高时, 应适当增加序列后面患者的预约服务时长, 减少序列前面的患者预约服务时长, 以实现患者的等待成本和医生的加班成本最小. 随着一个服务周期内预约时间长度的增加, 序列后面的患者预约服务时长应越长, 如图4所示.

下面比较是否考虑相关性的门诊预约调度. 不考虑服务时间相关性的门诊预约调度, 即预约调度模型(5)中仅已知服务时间的一阶矩、支撑集和平均绝对偏差的上界, 最优的预约服务时长随序列先减少随后

平稳,如图5所示.由图6可知,不考虑服务时间相关性的平均成本高于考虑服务时间相关性的平均成本,且相关性越强,平均成本相差越大,同时随着患者数量的增加,成本相差百分比也随着增加,其中,成本相差百分比=(不考虑相关性的平均成本/考虑相关性的平均成本)-1.

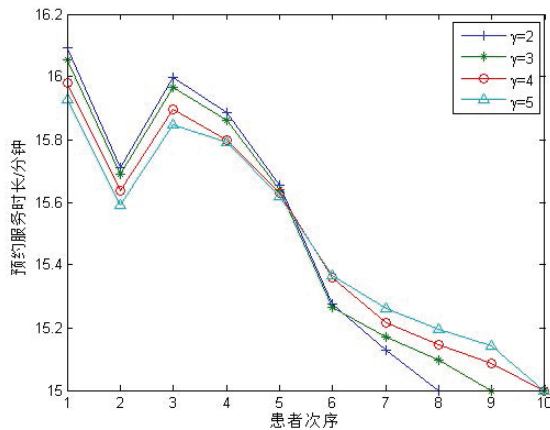


图3 不同 γ 下最优的预约服务时长

Fig. 3 The optimal appointment time allowances with different γ

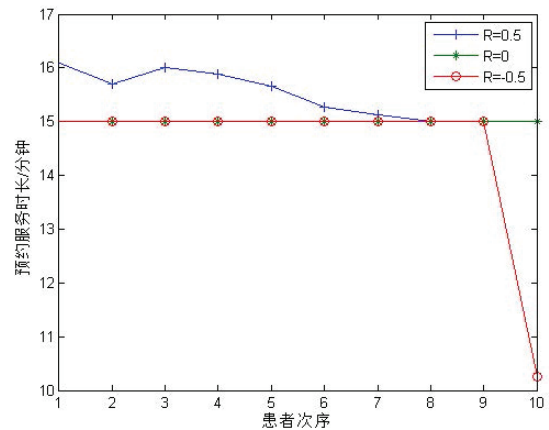


图4 不同 R 下最优的预约服务时长

Fig. 4 The optimal appointment time allowances with different R

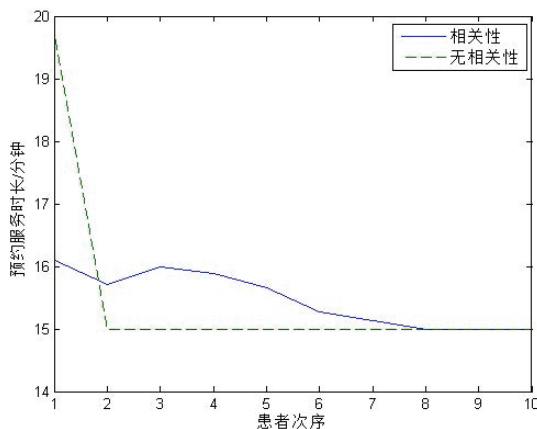


图5 无相关性与有相关性的预约服务时长比较

Fig. 5 The appointment time allowances comparison between no correlation and correlation

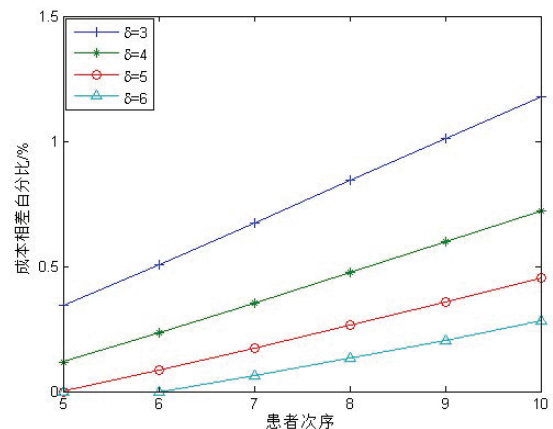


图6 无相关性与不同 δ 的成本相差百分比

Fig. 6 The percentage differences comparison between no correlation and different δ

文献[15]基于已知的服务时间概率分布,指出当服务时间独立同分布时,最优的预约服务时长呈先增后减的“圆顶型”,当服务时间不独立同分布时,则不具有该性质.本文研究了服务时间不独立的情况,发现最优的预约服务时长随序列呈递减的趋势,进一步丰富了门诊预约的调度策略.

3.2 不同类型患者的调度和排程

本节研究存在不同类型的患者的情况,确定最优的预约服务时长和服务次序.假设不同类型的患者服务时间的分布类型相同,参数不同;门诊预约系统通常存在不同类型的患者,主要包括复诊患者和初诊患者.不同类型的患者服务时间通常是不相同的,初诊患者服务时间的均值和标准差一般会高于复诊患者服务时间的均值和标准差,这主要是由于就诊前初诊患者的身体状况是未知的.本节的参数设置:复诊和初诊患者的人数比例2:3,分别编号1,2,...,n,按人数比例复诊患者编号在前,初诊患者编号在后,复诊患者服务时间的均值 μ_1 等于10 min,标准差 σ_1 等于2,初诊患者服务时间的均值 μ_2 等于15 min,标准差 σ_2 等于3,

$R = 0.5$, 变化一个周期内预约患者的人数 $n = 8 : 1 : 10$, $\delta = 1 : 1 : 7$, 医生与患者的相对时间成本 γ 的变化范围 $[0.5, 7]$.

表1 不同 δ 下的预约次序Table 1 The appointment sequences with different δ

δ	预约次序	δ	预约次序
无	(4,3,2,1,10,9,8,7,6,5)	4	(2,1,10,9,8,7,6,5,4,3)
1	(2,1,10,9,8,7,6,5,4,3)	5	(4,3,2,1,10,9,8,7,6,5)
2	(2,1,10,9,8,7,6,5,4,3)	6	(4,3,2,1,10,9,8,7,6,5)
3	(2,1,10,9,8,7,6,5,4,3)	7	(4,3,2,1,10,9,8,7,6,5)

表2 不同 γ 的预约次序Table 2 The appointment sequences with different γ

γ	预约次序	γ	预约次序
0.5	(2,1,9,8,7,6,5,4,3,10)	4	(2,1,10,9,8,7,6,5,4,3)
1	(2,1,9,8,7,6,5,4,3,10)	5	(10,9,8,7,6,5,4,3,2,1)
2	(2,1,10,9,8,7,6,5,4,3)	6	(10,9,8,7,6,5,4,3,2,1)
3	(2,1,10,9,8,7,6,5,4,3)	7	(10,9,8,7,6,5,4,3,2,1)

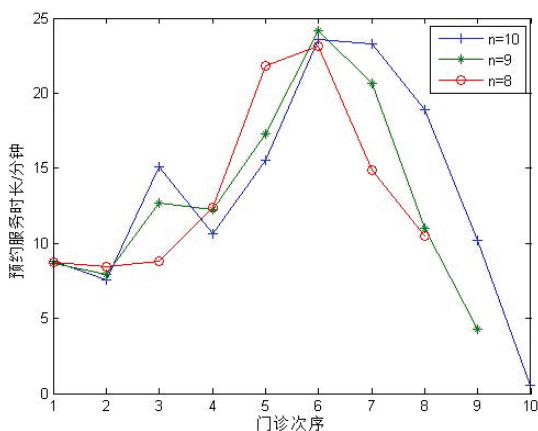
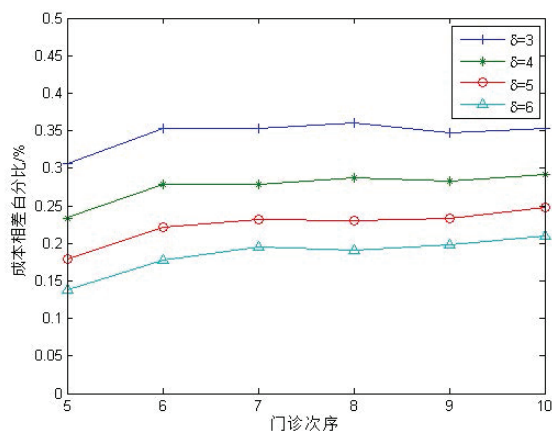


图7 不同类型患者及数量下的预约服务时长

Fig. 7 The appointment time allowances with different number and type of patients

图8 不同类型患者下无相关性与不同 δ 的成本相差百分比Fig. 8 The percentage differences comparison between no correlation and different δ with different patient types

当不考虑服务时间相关性时, 即预约排程模型 (18) 中仅已知服务时间的一阶矩、支撑集和平均绝对偏差的上界, 最优的预约服务次序把复诊患者安排在序列的前面, 初诊患者安排在序列的后面. 然而, 考虑服务时间相关性后, 相关性较强时, 最优的预约服务次序将复诊患者安排在序列的前面和后面, 初诊安排在序列的中间, 随着相关程度越来越小, 考虑服务时间相关性的排程策略和不考虑相关性的排程策略相同, 如表 1 所示. 存在不同就诊类型门诊患者的情况下, 最优的预约服务次序也与相对时间成本有关. 相对时间成本较高, 最优的预约服务次序把初诊患者安排在序列的前面, 复诊患者安排在序列的后面. 相对时间成本适中, 最优的预约服务次序把复诊患者安排在序列前面和后面, 初诊患者安排在中间. 相对时间成本较低, 最优的预约服务次序是复诊与初诊患者间隔排序, 具体见表 2. 医院管理者可根据实际情况, 选择最优的门诊预约服务次序.

当相对时间成本适中、相关性较强时, 由于将复诊患者安排在序列的前面和后面, 复诊患者的平均服务时间较短, 所以复诊患者的预约服务时长一般也比初诊患者的预约服务时长短, 可知此时的最优预约服务时长呈现先增后减的趋势, 如图 7 所示. 由图 8 可见, 同 3.1 节的结论相似, 在排程模型中, 不考虑服务时间相关性的平均成本要高于考虑服务时间相关性的平均成本.

4 结束语

本文针对门诊预约调度问题, 考虑不同类型的患者及服务时间的不确定性, 基于服务时间概率分布的支撑集和矩等部分信息, 利用绝对平均偏差刻画患者服务时间的相关性, 分别建立了分布式鲁棒优化门诊预

约调度和排程模型. 通过对偶理论及模型的数学性质进行等价变形, 将其化简成易处理的等价问题, 确定最优的预约服务时长和服务次序. 不仅在一定程度上避免用协方差矩阵刻画服务时间相关性造成的 NP 难问题, 提高了计算效率, 还能解决复诊和初诊患者的门诊排程问题. 数值分析结果显示, 当门诊的服务次序固定时, 同种类型患者的预约服务时长总体呈现出递减的趋势. 考虑初诊和复诊两种类型的患者, 最优的服务次序同医生与患者的相对时间成本和服务时间的相关程度有关. 当相对时间成本较高或服务时间相关程度较大时, 应适当增加序列后面的患者预约服务时长, 减少序列前面的患者预约服务时长.

本文考虑了不同的患者类型和服务时间的不确定性, 而在门诊预约调度中还存在爽约、当天到达等患者行为, 因此, 将考虑不同患者行为下的门诊预约调度和排程及相应的超订、预留机制作为后续的研究.

参考文献:

- [1] 罗 利, 石应康. 医疗服务资源调度优化理论、方法及应用. 北京: 科学出版社, 2014.
Luo L, Shi Y K. Theory, Method and Application of Healthcare Resource Scheduling. Beijing: Science Press, 2014. (in Chinese)
- [2] 曹萍萍, 唐加福, 姜博文. 多个医生环境下考虑取消预约的门诊预约策略. 系统工程学报, 2015, 30(5): 607–618.
Cao P P, Tang J F, Jiang B W. Outpatient appointment policy with multiple physicians considering cancellation of appointment. Journal of Systems Engineering, 2015, 30(5): 607–618. (in Chinese)
- [3] Begen M A, Queyranne M. Appointment scheduling with discrete random durations. Mathematics of Operations Research, 2011, 36(2): 240–257.
- [4] Ge D, Wan G, Wang Z, et al. A note on appointment scheduling with piecewise linear cost functions. Mathematics of Operations Research, 2013, 39(4): 1244–1251.
- [5] Castaing J, Cohn A, Denton B, et al. A stochastic programming approach to reduce patient wait times and overtime in an outpatient infusion center. IIE Transactions on Healthcare Systems Engineering, 2016, 6(3): 111–125.
- [6] 姜博文, 唐加福, 阎崇钧. 可加班门诊预约能力分配问题的期望方差模型. 系统工程学报, 2015, 30(2): 259–268.
Jiang B W, Tang J F, Yan C J. Expectation and variance model of capacity allocation in outpatient appointment considering overtime. Journal of Systems Engineering, 2015, 30(2): 259–268. (in Chinese)
- [7] Denton B, Viapiano J, Vogl A. Optimization of surgery sequencing and scheduling decisions under uncertainty. Health Care Management Science, 2007, 10(1): 13–24.
- [8] Mak H Y, Rong Y, Zhang J. Sequencing appointments for service systems using inventory approximations. Manufacturing & Service Operations Management, 2014, 16(2): 251–262.
- [9] Delage E, Ye Y. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Operations Research, 2010, 58(3): 595–612.
- [10] Gabrel V, Murat C, Thiele A. Recent advances in robust optimization: an overview. European Journal of Operational Research, 2014, 235(3): 471–483.
- [11] Hanasusanto G A, Roitch V, Kuhn D, et al. A distributionally robust perspective on uncertainty quantification and chance constrained programming. Mathematical Programming, 2015, 151(1): 35–62.
- [12] Kong Q, Lee C Y, Teo C P, et al. Scheduling arrivals to a stochastic service delivery system using copositive cones. Operations Research, 2013, 61(3): 711–726.
- [13] Mak H Y, Rong Y, Zhang J. Appointment scheduling with limited distributional information. Management Science, 2014, 61(2): 316–334.
- [14] Qi J. Mitigating delays and unfairness in appointment systems. Management Science, 2017, 63(2): 566–583.
- [15] Denton B, Gupta D. A sequential bounding approach for optimal appointment scheduling. IIE Transactions, 2003, 35(11): 1003–1016.

作者简介:

王珊珊(1991—), 女, 河北承德人, 博士生, 研究方向: 医疗运作管理, Email: shshwang.bit@163.com;

李金林(1955—), 男, 北京人, 博士, 教授, 博士生导师, 研究方向: 医疗与健康运作管理, 收益管理, Email: jinlinli@bit.edu.cn;

彭 春(1988—), 男, 山东临沂人, 博士生, 研究方向: 医疗资源配置优化, Email: pengchun12.18@163.com;

冉 伦(1977—), 男, 辽宁营口人, 博士, 教授, 博士生导师, 研究方向: 医疗与健康运作管理, Email: ranlun@bit.edu.cn.