阅读笔记:
Data-Driven Patient Scheduling in Emergency Departments: A Hybrid Robust-Stochastic Approach

## 一、名词解释:

- Door-to-provider上门就医
- Nested sets：嵌套集合。

a set containing a chain of subsets, forming a hierarchical structure, like Russian dolls.

来自 <https://en.wikipedia.org/wiki/Nested_set>

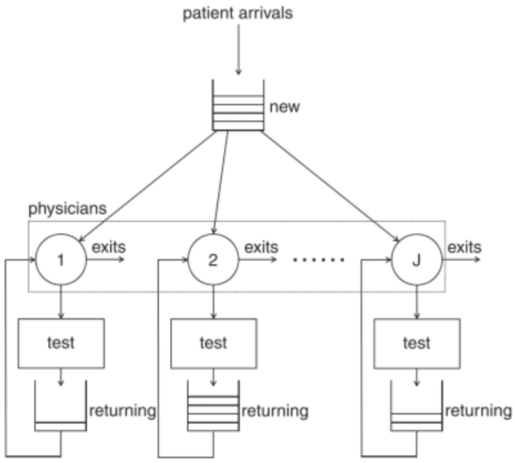
## 二、变量表:

| | Def | tips | 对应英文字母（仅参考） | |
|---|---|---|---|---|
| t | Time point | | | |
| $\mathscr{J}$ | The set of physicians | | j | |
| $\mathscr{N}$ | The set of new patients | | n | |
| $\mathscr{C}$ | the set of patients being seen | | c | |
| $\mathscr{R}$ | the set of returning patients | the dependence on $t$ is suppressed | r | |
| $\mathscr{C}_j$ | the set of patients being seen by physician $j$ | 1. $\mathscr{C}=\bigcup_{j\in\mathscr{J}}\mathscr{C}_j$<br>2. $\mathscr{C}_j=\emptyset$<br>IFF physician $j$ is available at time $t$<br>3.Otherwise, $\mathscr{C}_j$<br>has exactly one patient | | |
| $\mathscr{R}_j$ | the set of returning patients to be seen by physician $j$ | $\mathscr{R}=\bigcup_{j\in\mathscr{J}}\mathscr{R}_j$ | | |
| $\mathscr{W}$ | the set of patients in the waiting area | $\mathscr{W}=\mathscr{N}\cup\mathscr{R}$ | w | |
| $\mathscr{I}$ | the set of patients in the ED, excluding those sent to tests or treatments | $\mathscr{I}=\mathscr{W}\cup\mathscr{C}$ | i | |
| $\tilde{s}_{ij}$ | the consultation time of patient $i$ if he would be seen by physician $j$ | RV(see 四) | | |
| $F_{ij}$ | the cumulative distribution function of $\tilde{s}_{ij}$ | | | |
| $\mathscr{S}_{ij}$ | Set of $\tilde{s}_{ij}$ | $\mathscr{S}_{ij}=\{s_{ij}(1),\ldots,s_{ij}(N_{ij})\}$ | | |
| $\underline{s}_{ij}$ | smallest $\tilde{s}_{ij}$ | | | |
| $\bar{s}_{ij}$ | greatest $\tilde{s}_{ij}$ | | | |
| $\tilde{\mathbf{s}}$ | random vector of all the consultation times | $\tilde{\mathbf{s}}=(\tilde{s}_{ij})_{i\in\mathscr{I},j\in\mathscr{J}}$ | | |
| $\mathscr{S}$ | Product space | $\mathscr{S}=\prod_{i\in\mathscr{I},j\in\mathscr{J}}\mathscr{S}_{ij}$. | | |
| φ | The function to specify the assignment of waiting patients to physicians | $\varphi:\mathscr{W}\to\mathscr{J}$ | | |
| $\varphi(i)$ | The physician of patient i | | | |
| Φ | Function to specify sequencing decisions | $\Phi:\mathscr{W}\to\mathscr{P}(\mathscr{W})$, | | |
| $\mathscr{P}(\mathscr{W})$ | power set of $\mathscr{W}$ | | | |
| $\Phi(i)$ | the set of patients to be seen by the same physician before patient i | | | |
| $\mathscr{A}$ | the set of all admissible schedules | $(\varphi,\Phi)\in\mathscr{A}$ | | |
| $\tau_i$ | delay target | | | |
| π | Arrangement function | $\pi:\mathscr{S}\to\mathscr{A}$ | | |
| $\Psi$ | the set of all arrangements | | | |

Noted:

## 三、顾客流的排队网络模型

**Figure 1.** Queueing Network Model of Patient Flow in an ED




## 四、关于随机变量s_tilde_ij

$\tilde{s}_{ij}$

- Def: the consultation time of patient $i$ if he would be seen by physician $j$
- Remaining Consultation Time WHEN
  $i \in \mathscr{C}_j$
- $\{\tilde{s}_{ij} : i \in \mathscr{I}, j \in \mathscr{J}\}$
  a set of mutually independent random variables
- CDF:
  $F_{ij}$
  estimated using the records of physician j' s consultation times
- Choose empirical dist. in the implementation.
- Discrete random variable
  from a finite set
  $\mathscr{S}_{ij} = \{s_{ij}(1), \ldots, s_{ij}(N_{ij})\}$
- The smallest and greatest numbers
  $\underline{s}_{ij}$ and $\bar{s}_{ij}$

## 五、模型推演

Assignment function of patients:

$\varphi: \mathscr{W} \to \mathscr{J}$ 向physician派patients的函数.

Sequencing function:

Φ(i)为 所有和i看同一个医生且排在i之前的病人集合.

$$\varphi(k) = \varphi(i) \quad \text{for } k \in \Phi(i) \text{ and } i \in \mathscr{W}. \qquad (2)$$

同样地，如果我们有

$\varphi(i) = \varphi(k)$

代表i和k病人是同一个医生的病人，那么我们就会是i在k之前或者k在i之前，得到

$$\Phi(i) \subset \Phi(k) \quad \text{or} \quad \Phi(k) \subset \Phi(i). \qquad (3)$$

Admissible schedule: $\mathscr{A}$. 理解：不给任何一个patient安排医生的schedule是admissible吗

对于第k个病人，他的waiting time 可以写成:

$$w_k(\mathbf{s}, (\varphi, \Phi)) = \sum_{\ell \in \mathscr{C}_{\varphi(k)}} s_{\ell\varphi(k)} + \sum_{\ell \in \Phi(k)} s_{\ell\varphi(k)}, \qquad (4)$$

注意phi(k)为病人k的医生，第一个求和表示医生正在问诊的病人的剩余咨询时间，第二个求和表示排在病人k前面的病人的总咨询时间。

目标：最大化（病人问诊）联合概率

$$\begin{aligned} \max \quad & \mathbb{P}(w_i(\tilde{\mathbf{s}}, \pi(\tilde{\mathbf{s}})) \le \tau_i : i \in \mathscr{W}) \\ \text{s.t.} \quad & \pi \in \mathscr{V}. \end{aligned} \qquad (5)$$

A realization of V:

$$\begin{aligned} \max \quad & \mathbb{P}(w_i(\tilde{\mathbf{s}}, \pi(\tilde{\mathbf{s}})) \le \tau_i : i \in \mathscr{W}) \\ \text{s.t.} \quad & \pi \in \mathscr{V}_1. \end{aligned} \qquad (6)$$

棘手：递归程序带来的维数灾难

简化成static arrangements:

$\mathscr{V}_0 = \{\pi \in \mathscr{V} : \pi(\mathbf{s}_1) = \pi(\mathbf{s}_2) \text{ for } \mathbf{s}_1, \mathbf{s}_2 \in \mathscr{S}\}.$

全部consultation times对应相同(全体通用）的admissible arrangements，提取出这样的pai

【理解】丢掉了的那些按<1的概率admissible 的arrangement会有最优解吗

?

待理解:

$$\begin{aligned} \max \quad & \mathbb{P}(w_i(\tilde{\mathbf{s}}, \pi(\tilde{\mathbf{s}})) \le \tau_i : i \in \mathscr{W}) \\ \text{s.t.} \quad & \pi \in \mathscr{V}. \end{aligned} \qquad (5)$$

This formulation, however, cannot be implemented, because to determine the admissible schedule, one is required to know the realization of $\tilde{\mathbf{s}}$ in advance. To fix this issue, we should confine feasible solutions to (5) within the set of *nonanticipative* arrangements, which do not rely on future information to determine the patients to be seen when physicians become available. To specify a nonanticipative arrangement, we need to determine the assignment and sequencing decisions in a sequential manner. Let $w(1) \le w(2) \le \cdots$ be the times when

全部consultation times对应相同(全体通用）的admissible arrangements，提取出这样的pai
【理解】丢掉了的那些按<1的概率admissible 的arrangement会有最优解吗
?

相对于 $V$, $V_1$ 添加了future info. 等于增加约束，因此会是子集；在 $V_1$ 基础上，$V_0$ 则筛出静态 arrangement 的集合.

## 定理解析

**Proposition 1.** *Let $\mathcal{V}$, $\mathcal{V}_0$, and $\mathcal{V}_1$ be the sets of all arrangements, static arrangements, and nonanticipative arrangements, respectively. Then, $\mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}$.*

优化问题简化:

$$\begin{array}{ll} \max & \mathbb{P}(w_i(\tilde{\mathbf{s}}, \mu) \le \tau_i : i \in \mathcal{W}) \\ \text{s.t.} & \mu \in \mathcal{A}. \end{array} \tag{7}$$

Formulate 7 的 solution已经不会直接地rely on $\tilde{S}$.

## Hybrid robust stochastic approach

Consider the function $w_k$ given by (4) and extend its domain to $\mathbb{R}_+^{|\mathcal{I}||\mathcal{J}|} \times \mathcal{A}$. Under $\mu \in \mathcal{A}$, the set

$$\mathcal{X}(\mu) = \{x \in \mathbb{R}_+^{|\mathcal{I}||\mathcal{J}|} : w_k(x, \mu) \le \tau_k \text{ for all } k \in \mathcal{W}\}$$

is a convex polyhedron in $|\mathcal{I}| \cdot |\mathcal{J}|$ dimensions. Then, we may rewrite (7) as

$$\begin{array}{ll} \max & \mathbb{P}(\tilde{\mathbf{s}} \in \mathcal{X}(\mu)) \\ \text{s.t.} & \mu \in \mathcal{A}, \end{array} \tag{8}$$

### Hyperrectangular

$\mathcal{Z} = \prod_{i \in \mathcal{I}, j \in \mathcal{J}} [0, d_{ij}]$

个人理解：在凸多面体中间提取一个超矩形的子集

### 超矩形子集下的优化目标函数

$$\mathbb{P}(\tilde{\mathbf{s}} \in \mathcal{Z}) = \prod_{i \in \mathcal{I}, j \in \mathcal{J}} \mathbb{P}(0 \le \tilde{s}_{ij} \le d_{ij}) = \prod_{i \in \mathcal{I}, j \in \mathcal{J}} F_{ij}(d_{ij}),$$

这里d_ij表示一个正数

作者的假设：原问题的概率更大，hyperrectangular子集的概率更大；因此我们可以去寻找最优的Hyperrectangular子集问题的解用来作为近似最优，降低计算复杂度。

### 超矩形子集的族

$$\mathcal{H} = \left\{ \mathcal{S} \cap \prod_{i \in \mathcal{I}, j \in \mathcal{J}} [0, d_{ij}] : (d_{ij})_{i \in \mathcal{I}, j \in \mathcal{J}} \in \mathcal{S} \right\}.$$

可以看做s_tilde的不确定集的族

对于以上超矩形子集族的一个子集L

$$\begin{array}{lll} \max & \mathbb{P}(\tilde{\mathbf{s}} \in \mathcal{Q}) & \\ \text{s.t.} & w_k(\mathbf{s}, \pi(\mathbf{s})) \le \tau_k, & k \in \mathcal{W},\ \mathbf{s} \in \mathcal{Q} \\ & \mathcal{Q} \in \mathcal{H},\ \pi \in \mathcal{V}_1. & \end{array} \tag{9}$$

H看做不确定集，引入RO；目标函数中有RV，视为stochastic，所以本文是hybrid robust-stochastic approach.

### A simplified form of (9)

$$\begin{array}{lll} \max & \mathbb{P}(\tilde{\boldsymbol{\sigma}} \in \mathcal{B}) & \\ \text{s.t.} & \theta(\boldsymbol{\sigma}, \psi(\boldsymbol{\sigma})) \le \boldsymbol{\eta}, & \boldsymbol{\sigma} \in \mathcal{B} \\ & \mathcal{B} \in \mathcal{F},\ \psi \in \mathcal{G}. & \end{array} \tag{10}$$

Theorem 1 provides a simplified form of (10).

**Theorem 1.** *Let $(\mathbf{b}^*, \nu^*) \in \mathbb{R}^M \times \mathcal{D}$ be an optimal solution to the following problem:*

$$\begin{array}{lll} \max & \sum_{k=1}^{M} \ln G_k(b_k) & \\ \text{s.t.} & \theta(\mathbf{b}, \nu) \le \boldsymbol{\eta}, & \mathbf{b} = (b_1, \ldots, b_M) \\ & \mathbf{b} \in \mathbb{R}^M,\ \nu \in \mathcal{D}. & \end{array} \tag{11}$$

*Let $\mathcal{B}^*$ be the hyperrectangular subset in $\mathcal{F}$ with boundary value $\mathbf{b}^*$ and $\psi^*$ be the constant function with $\psi^*(\boldsymbol{\sigma}) = \nu^*$ for $\boldsymbol{\sigma} \in \mathbb{R}^M$. Then, $(\mathcal{B}^*, \psi^*)$ is an optimal solution to (10).*

注意psai_star是常值函数，也就是说consultation times的取值不会影响最优的arrangement，这里和第三章中的static对应。

另外theta是b的非减函数，boundary value满足约束意味着其他的取值也会满足。

这里作者的处理，static的提出，不同的consultation times对arrangement是无差异的，保证了最终解的可行性；hyperrectangular的假设，保证了目标函数联合概率的相互独立性，降低计算复杂度。

**Corollary 1.** *Let $(\mathbf{d}^*, \mu^*) \in \mathcal{S} \times \mathcal{A}$ be an optimal solution to the following problem:*

$$\begin{aligned} \max \quad & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \ln F_{ij}(d_{ij}) \\ \text{s.t.} \quad & w_k(\mathbf{d}, \mu) \le \tau_k, \qquad k \in \mathcal{W},\ \mathbf{d} = (d_{ij})_{i\in\mathcal{I}, j\in\mathcal{J}} \\ & \mathbf{d} \in \mathcal{S},\ \mu \in \mathcal{A}. \end{aligned} \tag{12}$$

*Let $\mathcal{Q}^*$ be the hyperrectangular uncertainty set in $\mathcal{H}$ with boundary value $\mathbf{d}^*$ and $\pi^*$ be the static arrangement with $\pi^*(\mathbf{s}) = \mu^*$ for $\mathbf{s} \in \mathcal{S}$. Then, $(\mathcal{Q}^*, \pi^*)$ is an optimal solution to (9).*

引理1其实是在static 的情况下对hyperrectangular set扩展到uncertainty set的情况

这里的RO结果取到boundary value的含义是最长的consultation times组合（worst case）作为优化的标准。