



## 《传染病动力学传播模型入门》教学改革实践与思考 附件材料

### 1 《传染病动力学传播模型入门》课程案例清单

#### 案例 1: HIV 传播动力学模型建模及应用

##### (1) Excel 入门练习（吸毒人群 HIV 感染干预模型）

基于课堂案例中注射吸毒人群的数量及人群特征，构建模型框架。Excel 表格提供模型初始人群个数及初始假设患病率，并对注射行为进行假设（每天注射 1 次），利用 Excel 计算每个模型步长下仓室人数的变化，并对 HIV 患病率、艾滋病致死人数等指标进行估计。

##### (2) MATLAB 建模及应用（男同人群行为干预 HIV 模型）

基于课堂案例讲解的男同人群的 HIV 模型，根据教学团队所提供的 MATLAB 代码运行模型，汇报 HIV 新发感染人数、HIV 死亡人数等指标在内的模型结果。进而分小组讨论，按照课堂要求，修改模型中的部分参数或模型结构，修改对应代码，回答教师问题，并总结实操过程中的疑难问题向教师提问。

#### 案例 2: COVID-19 防控措施效果评估的传播动力学模型建模及应用

##### (1) COVID-19 评估模型的构建

基于某市的新冠疫情数据，思考构建以 SEIR 框架为基础的 COVID-19 防控措施效果评估的仓室模型，确定模型所包含的仓室结构及对应的微分方程组。

##### (2) COVID-19 评估模型的应用

分小组讨论确定组内的模型结构，并基于构建的 COVID-19 评估模型，撰写 MATLAB 代码，分小组讨论并总结汇报模型的评估结果，实践操作期间可随时向同学或教师提问。

#### 案例 3: 再生系数估计



基于某地每天新增的 COVID-19 确诊病例数，以及文献中指定代际时间分布，在 R 软件中，实现基本再生数  $R_0$  的计算。按照课堂要求，小组探索采用指数增长法、极大似然法、序贯贝叶斯、时间相关基本再生数等多种方法再生系数估计结果的差异。

#### 案例 4：人群接触模式调查与量化

##### （1）接触模式调查设计

了解人群接触模式调查的常用调查方式，以及以接触日志形式开展的调查所需的信息要素，分小组尝试设计接触模式调查问卷。

##### （2）R 语言人群接触模式量化及可视化

基于某市人群接触模式的调查数据，在 R 中整理计算得到人群接触模式的估计结果并进行可视化绘图。按照课堂要求，小组尝试按不同人群划分依据探索不同人群间接触模式，总结其中规律并汇报展示。

##### （3）延伸拓展——接触模式在传播动力学模型的应用

分小组探索获得的接触模式如何影响传播动力学模型的预测结果并汇报展示。



## 2 《传染病动力学传播模型入门》实习案例说明文档

### 2.1 HIV 传播动力学建模练习 (Practice 1)

#### 1、研究目的

建立一个简单的数学模型来描述 HIV 在某国注射吸毒者 (People who inject drugs, PWID) 中的传播情况。预测基线情景下, 该国 PWID 人群在未来 10 年的:

- ①HIV 患病率;
- ②因过量注射致死的死亡人数;
- ③HIV 相关死亡人数。

应用模型进行不同干预措施的效果评估: 比较三种干预措施下 (A. 美沙酮维持治疗; B. 清洁针具项目; C. 扩大 HIV 检测及 ART 治疗) 的 HIV 患病率, 及因注射吸毒和 HIV 分别造成的总生命损失 (total lives lost)。比较哪种干预最有效? 哪种干预措施可以最大比例预防 HIV 感染? 哪种干预措施总体上可以最大比例预防死亡?

#### 2、研究假设

(1) 注射吸毒人群可能处于以下三种状态之一: 未感染 (Uninfected)、感染 HIV 但无症状 (living with HIV, asymptomatic) 或出现艾滋病感染症状 (living with AIDS);

(2) 由于用药过量和暴力等因素 (与注射吸毒相关的常见健康风险), 注射吸毒人群的背景死亡率非常高;

(3) HIV 感染者可以在无症状期或 AIDS 期接受治疗。进入治疗阶段需要我们对感染者接受检测 and 治疗的比率做出假设;

(4) 处于无症状期和 AIDS 期的感染者因感染 HIV 具有额外的死亡率, 具体的死亡率取决于感染阶段;

(5) 进行 ART 治疗可以降低其传染性, 延长 HIV 感染者的无症状期, 并降低 HIV 引起的死亡率, 但这并不影响其他行为。

#### 3、建模步骤

- (1) 确定模型结构: 构建的模型包含了 5 个仓室, 模型示意图如下:

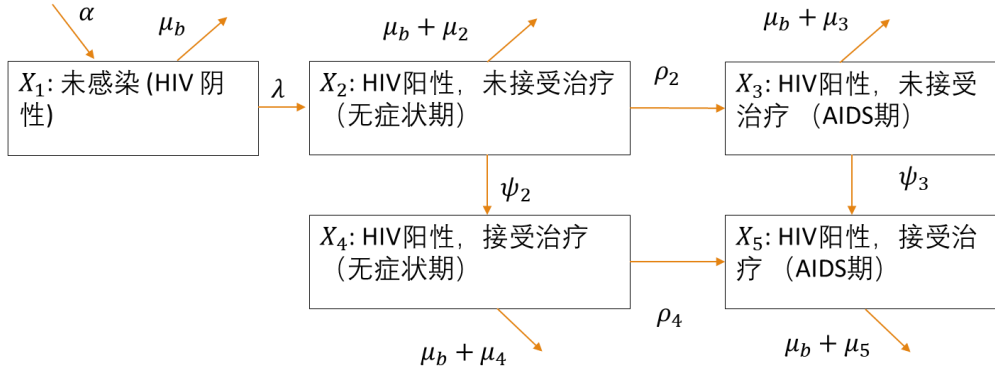


图 1 模型结构示意图

(2) 构建微分方程：

$$\frac{dX_1}{dt} = \alpha - \mu_b X_1 - \lambda X_1$$

$$\frac{dX_2}{dt} = \lambda X_1 - (\mu_b + \mu_2) X_2 - \psi_2 X_2 - \rho_2 X_2$$

$$\frac{dX_3}{dt} = \rho_2 X_2 - (\mu_b + \mu_3) X_3 - \psi_3 X_3$$

$$\frac{dX_4}{dt} = \psi_2 X_2 - (\mu_b + \mu_4) X_4 - \rho_4 X_4$$

$$\frac{dX_5}{dt} = \psi_3 X_3 + \rho_4 X_4 - (\mu_b + \mu_5) X_5$$

修正人口数量，使之与死亡人数相匹配（为简化）：

$$\alpha = \mu_b X_1 + (\mu_b + \mu_2) X_2 + (\mu_b + \mu_3) X_3 + (\mu_b + \mu_4) X_4 + (\mu_b + \mu_5) X_5$$

定义总人口：

$$X = X_1 + X_2 + X_3 + X_4 + X_5 = \sum_{i=1}^5 X_i$$

(3) 传染力计算：

$$\begin{aligned} \lambda &= np_s \left( \beta_2 \frac{X_2}{X} + \beta_3 \frac{X_3}{X} + \beta_4 \frac{X_4}{X} + \beta_5 \frac{X_5}{X} \right) = \frac{np_s}{X} \sum_{i=2}^5 \beta_i X_i \\ \lambda &= \left[ 1 - (1 - \beta_2)^{np_s \frac{X_2}{X}} \right] + \left[ 1 - (1 - \beta_3)^{np_s \frac{X_3}{X}} \right] + \left[ 1 - (1 - \beta_4)^{np_s \frac{X_4}{X}} \right] \\ &\quad + \left[ 1 - (1 - \beta_5)^{np_s \frac{X_5}{X}} \right] \\ &= \sum_{i=2}^5 \left[ 1 - (1 - \beta_i)^{np_s \frac{X_i}{X}} \right] \end{aligned}$$



## (4) 获取模型参数:

符号	取值	备注	参考文献
$\alpha$	根据死亡总人数计算		
$\mu_b$	0.0016	孟加拉国成人（35-39岁）的月死亡率，PWID的SMR为16	PWID死亡率数据来源于Mathers et al <sup>[5]</sup> . 孟加拉国死亡率数据来源于SEARO <sup>[6]</sup>
$\mu_2$	0.02	假设和其他人群一样	Li et al. <sup>[7]</sup>
$\mu_3$	0.22		Li et al. <sup>[7]</sup>
$\mu_4$	0.02		Li et al. <sup>[7]</sup>
$\mu_5$	0.08		Li et al. <sup>[7]</sup>
$\beta_1$	0.0087	单次注射的传染性（HIV）	PHAC <sup>[8]</sup>
$\beta_2$	0.02	单次注射的传染性（AIDS）	PHAC <sup>[8]</sup>
$\rho_2$	0.012		Li et al. <sup>[7]</sup>
$\rho_4$	0.0028		Assumed
$\psi_2$	0.0155		Li et al. <sup>[7]</sup>
$\psi_3$	0.0263		假定的
$n$	30	1个月的注射次数	假设每天注射一次
$p_s$	0.65	假设 2/3 的人共用针头	Azim et al. <sup>[1]</sup>

## 2.1.1 Excel 入门练习

## 1、文件说明

文件夹中“用 Excel 构建仓室模型练习 Solution base.xlsx”文件为练习数据；“用 Excel 构建仓室模型练习答案-Solution full.xlsx”为配套练习答案。

## 2、操作步骤

## ①基线模型构建步骤:

- 1) 利用 T3 到 T7 单元格的参数计算各人口组的死亡人数，记录到 P4 和 Q4；
- 2) 计算新进入人群的总人数（ $\alpha$ ，N4 单元格）以平衡总死亡人数；
- 3) 使用在问题 5 中得到的公式计算  $\lambda$ （M4 单元格）；
- 4) 使用  $\alpha$  和第 T 列的所有参数，计算各仓室的人数变化并记录，用 H4 到 L4 单元格来存储每个时间段各仓室人数变化；
- 5) 使用这些率更新次月的群体值（C5 到 G5 单元格）；
- 6) 将人口数单元格复制粘贴到第 12 个月，并将率的单元格复制粘贴到第 11 个月；
- 7) 确定患病率的计算公式（第 R 列），并计算到第 12 个月；



8) 同时可以尝试使用 X2 变化率方程的组成部分和高危人群来估计 HIV 发病率（第 O 列）。

②将模型向前运行 10 年，以获得 HIV 患病率以及因艾滋病或吸毒死亡的总人数的最终估计值。

③复制基线模型表格 3 次，并考虑三种可能的政策方案，修改对应的参数值

## 2.1.2 MATLAB 练习

### 1、文件说明

maincode-主程序：初始化、初始状态赋值、参数赋值、调用 controlFunction 函数、结果输出、画图；

controlFunction 函数：输出在某组参数下，调用 diffCalculator 函数，存储所有的流行病学效果指标；

diffCalculator 函数：输出某组参数下，根据传播力  $\lambda$ ，单个时间步长每个仓室的变化值及 incidence、HIV mortality、Total mortality；

### 2、操作步骤

运行 maincode 主程序即可得到模型结果。



## 2.2 COVID-19 动力学建模练习 (Practice 2)

### 1、案例背景

随着新型冠状病毒肺炎防控方案（第九版）和中央“20条”要求颁布，某市卫生部门拟快速预测管控措施开放后该市 COVID-19 流行趋势，评估当地卫生医疗系统的承受能力，故基于传播动力学模型对当地疫情流行趋势开展评估。

### 2、建模步骤

(1) 确定模型结构：构建了 SEIR (Susceptible-Exposed-Infected-Recovered) 模型，并在普通 SEIR 模型中，将 I 细分无症状感染者 (Asymptomatic)、有症状感染者 (Symptomatic) 2 大类，其中有症状感染者分为轻症 (Minor)、普通型 (General)、重症 (Severe)、危重症 (Critical) 4 类。考虑只有在危重症病例中才会发生死亡，感染者可通过治疗或自愈恢复到 R 仓。模型输出结果包括累计感染病例数、每日现存病例数、每日新增病例数、累计重症及危重症病例数、每日死亡病例数 5 部分。模型示意图如下：

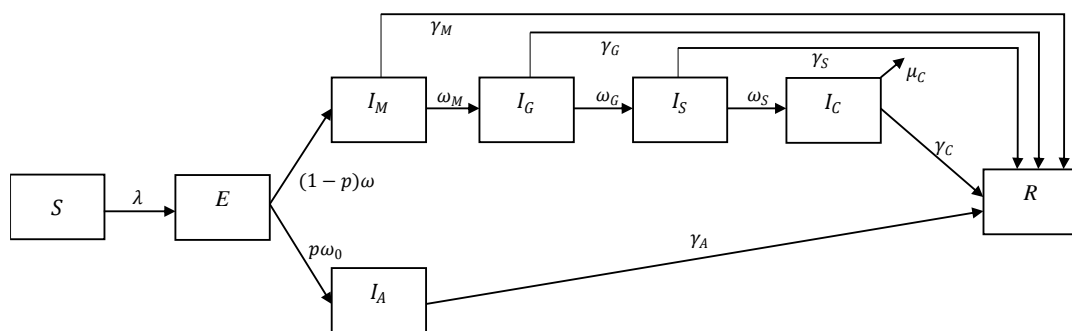


图 2 模型结构示意图

模型有以下重要假设，暂不考虑流入/新生人口、潜伏期和无症状感染者也具备一定传染性、重症和危重症患者因住院而阻断传播、仅处于危重症阶段的患者可能发生死亡，其余感染者可治愈或自愈恢复

(2) 构建微分方程：详见 Matlab 文件 diffCalculator。

(3) 收集模型参数：参数含义取值详见 Excel 文件和 Matlab 代码。

参数	释义
$\lambda$	传染力
$p$	无症状感染者的比例
$\omega_0/\omega$	潜伏期时长
$\omega_M \backslash \omega_G \backslash \omega_S \backslash \omega_C$	进展比例和进展速度的综合参数



---

$\gamma_{M\backslash G\backslash S\backslash C}$	考虑恢复比例和时长的综合参数
$\mu$	死亡率

(4) 模型代码实现：详见 Matlab 代码文件：主要包括三个程序，程序 nCoV 为主程序，主要为参数输入为结果输出；程序 diffCalculator 定义微分方程；程序 controlFunction 定义传染力计算和结果计算规则。

(5) 结果展示，使用 Excel、R 等软件即可完成绘图。





## 2.3 再生系数 $R_0$ 的估计 (Practice 3)

### 1、数据准备

根据“cases01.xlsx”中某地每天新增的 COVID-19 确诊病例数（共 27 条记录），将数据导入 R 软件并整理为所需格式（参考  $R_0$  package 说明）。

- 提示一：数据应至少包含：

- ①发病数（incidence data），可以是 vector、data.frame、list 格式；

- ②时间向量，时间间隔（time step）一般默认为每天。

- 提示二：导入并整理为所需数据格式后，可以使用 check.incid

### 2、参数依据

参考中国疾病预防控制中心 2020 年发表在新英格兰医学杂志的文章结果<sup>[1]</sup>，指定代际时间（Generation time, GT）分布（此处假设代际时间已知）。

截取内容如下：

Results: Among the first 425 patients with confirmed NCIP, the median age was 59 years and 56% were male. The majority of cases (55%) with onset before January 1, 2020, were linked to the Huanan Seafood Wholesale Market, as compared with 8.6% of the subsequent cases. The mean incubation period was 5.2 days (95% confidence interval [CI], 4.1 to 7.0), with the 95th percentile of the distribution at 12.5 days.

### 3、实现（基于 R）

基于“ $R_0$ ” package

R 语法参考：

```
generation.time(type = c("empirical", "gamma", "weibull", "lognormal"),
                val = NULL,
                step = 1)
```

**Task:** 采用指数增长法(exponential growth, EG)、极大似然法(maximum likelihood estimation, ML)、序贯贝叶斯方法(sequential Bayesian method, SB)、时间相关基本再生数(time dependent reproduction numbers, TD) 计算基本再生数  $R_0$

\* 提示：这些方法输出包含置信区间

#### ①EG 法（指数增长法）

```
est.R0.EG (epid, GT, t = NULL, begin = NULL, end = NULL, time.step = 1)
```

#### ②ML 法（极大似然法）

```
est.R0.ML (epid, GT, t = NULL, begin = NULL, end = NULL, time.step = 1)
```

#### ③SB 法（序贯贝叶斯方法）

```
est.R0.SB (epid, GT, t = NULL, begin = NULL, end = NULL, time.step = 1)
```



④TD 法（时间相关基本再生数）

`est.R0.TD (epid, GT, t = NULL, begin = NULL, end = NULL, time.step = 1)`

**\* 注：** SB 法和 TD 法需进一步计算均值，

另外可用 `plot()` 展示估计结果和 `plotfit()` 展示模型估计结果与原始流行曲线的拟合效果

参考文献：

- [1] Li Q, Guan XH, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia [J]. N Engl J Med. 2020;382(13):1199-1207. doi:10.1056/NEJMoa2001316



## 2.4 Contact Pattern 练习 (Practice 4)

### 1、数据准备

文件为 2023 年年初某周进行的人群接触模式调查的数据，旨在针对人群一日内的非家庭接触构建最为常见的分年龄的人群接触矩阵。

数据应至少包含：①受访者 ID（计算各年龄段的调查人数）；②受访者年龄/年龄段；③被接触者年龄/年龄段

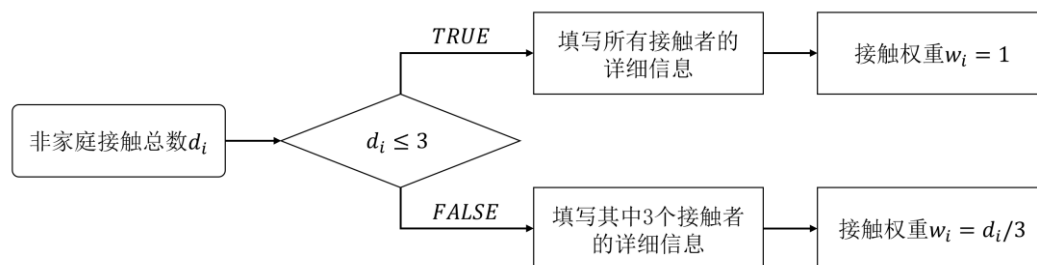
### 2、数据加权处理<sup>1</sup>

#### ① 受访者权重

用 `autumn` 包（采用 *Iterative proportional fitting* 方法）以及该地区的人口结构数据（可由统计年鉴等官方数据库获得）对受访者进行加权，以平衡样本的代表性与均衡性。

#### ② 接触行为权重

非家庭接触（非共同居者）的接触行为权重需根据接触人数分情况处理，报告的接触人数大于 3 时，对未填写具体年龄的其余被接触者的信息进行推断。加权过程如下图：



若  $d_i=6$ ，3 个接触者年龄分别为 18-25、26-30、30-35 岁，则  $w_i=6/3=2$ ，假设其分别接触 18-25、26-30、30-35 岁年龄段的人数各 2 人；若年龄分别为 18-25、18-25、26-30 岁，假设其分别接触 18-25 岁的人数为 4，26-30 岁的人数为 2。

### 3、样本接触矩阵A的计算

将所调查的样本按受访者及其接触对象划分为 14 个年龄层（0-4、5-9、10-

<sup>1</sup> 非必须操作，本练习的模拟数据已经过加权处理



14、15-19、20-24、25-29、30-34、35-39、40-44、45-49、50-54、55-59、60-64、 $\geq 65$ ), 由于调查在 15 岁以上人群中进行, 因此将生成  $11 \times 14$  的样本接触矩阵  $A$ 。

在计算样本接触矩阵  $A$  时, 首先定位每个受访者的年龄组  $i$  (由下向上第  $i$  行), 然后将其接触人员填充至对应的列。所有数据对填充后, 将  $A$  的每行除以该组的受访者总数, 即得到平均每人的接触人数。

矩阵的每一个元素  $a_{ij}$  代表年龄组  $i$  受访者与报告年龄组  $j$  接触者之间的每日接触的平均次数。例如:  $a_{11}$  为第一个年龄组的 1 个受访者与第一个年龄组的平均接触人数; 如  $a_{12}$  为第 1 个年龄组的 1 个受访者与第 2 个年龄组的平均接触人数。

$$A = \begin{matrix} & \begin{matrix} j \\ a_{n1} & a_{n2} & a_{n3} & \cdots & \cdots & \cdots & a_{nn} \end{matrix} \\ \begin{matrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ a_{31} & a_{32} & a_{33} & \cdots & \cdots & \cdots & a_{3n} \\ a_{21} & a_{22} & a_{23} & \cdots & \cdots & \cdots & a_{2n} \\ a_{11} & a_{12} & a_{13} & \cdots & \cdots & \cdots & a_{1n} \end{matrix} & \begin{bmatrix} & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{bmatrix} & \\ & \begin{matrix} i \end{matrix} \end{matrix}$$

#### 4、人群接触矩阵 $C$ 的计算

由于人群接触矩阵具有对称性 ( $c_{ij} \cdot N_i = c_{ji} \cdot N_j$ , 即两个年龄段之间的接触总数相等), 此练习中将采用①简单平均法<sup>[1]</sup>; ②最小二乘法 (LS)<sup>[2]</sup>, 根据矩阵  $A$  获得真实接触矩阵  $C$ 。故有:

##### ① 简单平均法<sup>2</sup>:

<sup>2</sup> 由于调查在 15 岁以上人群中进行, 无 0-4 岁、5-9 岁、10-14 岁年龄组的受访者, 用简单平均法时



$$\hat{c}_{ij} = \frac{a_{ij}N_i + a_{ji}N_j}{2N_i}, \quad \forall i \neq j$$

$$\hat{c}_{ii} = a_{ii}$$

②最小二乘法 (LS) :

$$\begin{aligned} \hat{c}_{ij} &= \underset{c_{ij}}{\operatorname{argmin}} (c_{ij} - a_{ij})^2 + \left[ \left( \frac{N_i}{N_j} \right) c_{ij} - a_{ji} \right]^2 \\ &= \underset{c_{ij}}{\operatorname{argmin}} \left( 1 + \frac{N_i^2}{N_j^2} \right) c_{ij}^2 - \left( 2a_{ij} + 2\frac{N_i}{N_j} a_{ji} \right) c_{ij} + (a_{ij}^2 + a_{ji}^2) \\ &= \frac{a_{ij} + \frac{N_i}{N_j} a_{ji}}{1 + \frac{N_i^2}{N_j^2}}, \quad \forall i, j, i < j \\ \hat{c}_{ji} &= \frac{N_i}{N_j} c_{ij}, \quad \forall i, j, i < j \end{aligned}$$

特别地,

$$\begin{aligned} \hat{c}_{ii} &= \underset{c_{ij}}{\operatorname{argmin}} (c_{ii} - a_{ii})^2 \\ &= a_{ii}, \quad i = 1, 2, 3, \dots, n \end{aligned}$$

参考文献:

- [1] Feehan DM, Mahmud AS. Quantifying population contact patterns in the United States during the COVID-19 pandemic. Nat Commun. 2021;12(1):893.
- [2] Guo X, Liu Z, Yang S, et al. Contact pattern, current immune barrier, and pathogen virulence determines the optimal strategy of further vaccination. Infect Dis Model. 2023;8(1):192-202.

---

对于  $i = 1, 2, 3$  需使用另外的公式进行修正。

**附：练习数据说明**

本练习分别提供了 X 市 2023 年年初某周调查得到的非家庭接触模拟数据 (contact-sc.xlsx)，该数据是通过从加权处理后的模拟调查数据中提取受访者报告的非家庭接触信息整理得到的，数据形式如下：

ID	res_age	res_gender	nhh_cont	w_res	w_cont	cont_age	cont_gender
1	30-34	女	2	1.413	1	25-29	女
1	30-34	女	2	1.413	1	25-29	男
2	20-24	女	29	0.7291	9.6667	45-49	男
2	20-24	女	29	0.7291	9.6667	35-39	女
2	20-24	女	29	0.7291	9.6667	30-34	男
3	20-24	男	29	0.5351	9.6667	25-29	女
3	20-24	男	29	0.5351	9.6667	25-29	男
3	20-24	男	29	0.5351	9.6667	10-14	男
4	25-29	男	1	0.6665	1	25-29	男
5	35-39	男	10	0.4747	3.3333	55-59	女
5	35-39	男	10	0.4747	3.3333	55-59	男
5	35-39	男	10	0.4747	3.3333	≥65	女
6	55-59	女	2	2.8094	1	60-64	男
6	55-59	女	2	2.8094	1	≥65	男

**(1) 变量说明：**

ID：受访者的唯一标识符；

res\_age、res\_gender：受访者的年龄、性别；

cont\_age、cont\_gender：受访者报告的接触对象的年龄、性别；

nhh\_con：该受访者报告的非家庭的接触数（topcode：29）；

w\_res：受访者权重；

w\_cont：接触行为权重。

**(2) 数据阅读方式：**

表中每一横行表示一条详细的接触信息，下图中数据表中第 1、2 行表示由 1 号受访者报告的 2 条非家庭接触的详细信息。

ID	res_age	res_gender	nhh_cont	w_res	w_cont	cont_age	cont_gender
1	30-34	女	2	1.413	1	25-29	女
1	30-34	女	2	1.413	1	25-29	男
2	20-24	女	29	0.7291	9.6667	45-49	男



### (3) 代码说明:

`contact_matrix.R` 给出了运算接触矩阵的详细代码。代码的第二板块中 `computeContactDataMatrix` 函数给出了样本接触矩阵 $A$ 的计算方式；第三板块中 `estimateCM_Average` 和 `estimateCM_LSE` 函数分别提供了使用简单平均法和最小二乘估计法根据矩阵 $A$ 获得真实接触矩阵 $C$ 的过程；同时还提供了绘制接触矩阵热图的函数 `matrixplot` 函数。运行程序即可得到 Age-based 的接触模式图示结果。