

# 2018 年江西省研究生数学建模竞赛

参赛队号： 2018630

---

题目： 2284 组过敏原检测结果分析

---

——从过敏性疾病到减少 CO<sub>2</sub> 排放

---

## 2284 组过敏原检测结果分析

### ——从过敏性疾病到减少 CO<sub>2</sub> 排放

#### 摘要

本文针对过敏性疾病与时间、季节、性别、年龄、过敏原的相关性问题，将  $\chi^2$  检验与相关性示意图结合，分析各变量间有无关系、有何种程度关系。建立了 SVM 模型和贝叶斯网络模型并使用 SPSS Modeler14.1 进行实现，该模型可根据患者就诊信息推测所患过敏性疾病类型，准确率高达 93%；定量分析过敏原检测结果随时间变化趋势，结合过敏原数量、传播方式、其他影响因素综合考虑，选取 CO<sub>2</sub> 作为环保指标进行相关性分析。结合全文分析结果，对于减少 CO<sub>2</sub> 排放进行了倡议，并给出了具体的可行性措施。

对问题一，我们采用“总-分”的研究思路。使用 SPSS Statistics24.0 进行  $\chi^2$  检验，分别判断过敏性疾病与时间、季节、性别、年龄等变量间有无相关性。对于存在相关性的变量，进一步使用相关性示意图对不同类型的过敏性疾病与该变量间相关程度进行分析。

对问题二，通过建立 SVM 模型求得了时间、季节、性别、年龄、过敏原等变量对于过敏性疾病的重要性，即诱发权重。建立了贝叶斯网络模型以求得变量之间的相关性。使用 SPSS Modeler14.1 实现后，本模型可根据患者的就诊数据对其过敏性疾病类型进行推测，准确率高达 93.35%。因此，本模型求得的各变量对于不同类型过敏性疾病的重要性，即为不同类型疾病的病人过敏原检测结果之间的差异。

对问题三，定量分析 2013-2017 年过敏原检测结果变化趋势，结合过敏原数量、传播方式、其他影响因素（如政策等）综合考虑，选取 CO<sub>2</sub> 浓度作为待分析的环境保护指标，从中国气象局获取 2013-2017 年大气中 CO<sub>2</sub> 浓度数据并作出其变化趋势并与过敏原检测结果变化趋势进行对比分析，以找出两者之间的相关性。

对问题四，结合本文分析结果，从引发全球变暖、加剧花粉过敏等方面阐述了 CO<sub>2</sub> 对日常生活带来的影响。倡议市民在日常生活中通过一些细节来减少 CO<sub>2</sub> 排放以缓解呼吸道过敏性疾病的频繁，同时针对夏季“烧烤热”给出了具体的可行性措施。

最后，我们对模型的优缺点进行了评价，并提出了模型的改进方向和推广展望。

**关键词：**过敏性疾病； $\chi^2$  检验；SVM 模型；贝叶斯网络模型；CO<sub>2</sub>

## 一、问题重述

十九大报告中提出‘要像对待生命一样对待生态环境’，保护好生态环境就是保护未来。生态环境受到一定程度的污染，例如大气中的微量物质达到一定的浓度，可能对人类、动物等产生不利影响或危害，会导致各类过敏性疾病的发生，譬如皮肤过敏、呼吸道过敏、消化道过敏等。

关于生态环境污染与人体健康之间的关系国内外学者已有大量的研究。不同的环境指标和污染程度对人体健康的影响程度也不同，不同人体的各个系统的差异性导致其在日常生活中对接触物、吸食物具有不同的敏感性。Boezen H M<sup>[1]</sup>通过调查不同区域年龄在 50-70 岁之间的成年人发现，成人呼吸系统疾病症状的发生确实与大气污染有关。马关培<sup>[2]</sup>等人通过分析生态环境污染急性暴露对医院每天门诊人数之间的关系发现 PM<sub>2.5</sub> 和 NO<sub>2</sub> 的浓度短时间内升高会导致呼吸道过敏的就诊人数增加。

现给出某医院过敏原检测数据包括检测日期，性别，年龄，科室及各过敏原结果（包括普通豚草，艾蒿，尘螨组合，屋尘螨，猫毛等 17 种过敏原）呈阴性或阳性及其敏感性程度。

需要考虑如下问题：

- （1）试分析过敏性疾病与时间，季节，性别，年龄分别有什么样的关系；
- （2）结合第一问的相关影响因素，分析不同类型疾病的病人过敏原检测结果之间的差异；
- （3）定量分析从 2013 年到 2017 年过敏原检测结果的变化趋势，选取一个环境保护的指标（譬如 PM<sub>2.5</sub> 浓度），从网上查找相关数据分析两者之间的相关性；
- （4）根据上述分析结果，写一份生态环境保护的倡议书。

## 二、问题分析

### 2.1 问题一的分析

针对问题一，要求分析过敏性疾病与性别、年龄、季节、时间之间的关系。本问基于对数据的预处理，采用“总一分”的研究思路，首先通过 $\chi^2$ 独立性检验对过敏性疾病与时间、季节、性别、年龄之间是否存在显著性关系进行分析，当 $P < 0.05$ 时，认为差异具有统计学意义；然后通过对对应分析研究皮肤过敏、呼吸道过敏、消化道过敏三大过敏性疾病与性别、年龄、季节、时间之间存在什么样联系，并通过三大过敏性疾病与性别、年龄、季节、时间的对应分析因子负荷图和相关性分析图两者比较得出性别与过敏之间是否存在联系、哪个年龄段的人群较易过敏、哪个季节容易引起过敏以及随着时间的变化患过敏性疾病人群如何变化，并寻找可能的缘由做出解释。

### 2.2 问题二的分析

根据题意可知，问题二要求结合第一问的相关影响因素，分析不同类型疾病的病人过敏原检测结果之间的差异。为了更好的寻找题设所求的差异，我们借助 SPSS Modeler 软件建立 SVM 模型对过敏性疾病的相关影响因素进行分类，这种分类方法可以输出模型中预测变量的重要性，通过对其结果进行分析即可得到题设所求。接着结合第一问我们的分析结果与上述算法得到的预测变量的重要性，选用贝叶斯网络模型对重要的变量和第一问分析的结果进行组合分析，从而得到重要的过敏原与季节、性别、年龄之间可能的联系。

### 2.3 问题三的分析

据题意可知，问题三要求定量分析从 2013-2017 年过敏原检测结果的变化趋势，选取一个环境保护指标并分析两者之间的相关性。首先定量分析 2013-2017 年内 17 种过敏原阳性率，作出其变化趋势示意图并分析其变化趋势，结合过敏原数量、传播方式、其他影响因素（如政策等）综合考虑选取环境保护指标，然后对选取的主要环境保护指标在 2013-2017 年的变化趋势进行分析，最后对两者之间的相关性进行评价。

### 2.4 问题四的分析

据题意可知，问题四需要结合本题分析结果，结合生态环境保护的相关知识、第三问选取的环保指标、第一、二问关于各变量与过敏性疾病间的相关性进行分析，并给出针对、有效的倡议。

### 三、模型假设

- 1、题中所给内容和数据资料都是真实可信的；
- 2、除前往体检、体检中心进行过敏原检测的人员外，假设所有前往医院就诊并进行过敏原检测的患者都存在过敏性疾病或与其类似的临床表征；
- 3、数据中过敏原检测结果全为 0（过敏性疾病检测结果为阴性）的个体，假设其所患疾病为表征与过敏性疾病相似的其他类疾病，即未患过敏性疾病；
- 4、数据中过敏原检测结果存在不为 0（过敏性疾病检测结果为阳性）的个体，假设其过敏原只为所给数据中的过敏原，不考虑其他过敏原的影响；
- 5、假设各过敏原之间相互独立，暂时不考虑各过敏原之间相互作用对人体产生影响；
- 6、假设每位到医院就诊的过敏性疾病患者前往的科室为其过敏类型临床表征最明显的科室，不考虑多种过敏性疾病同时表征的情况；

### 四、符号说明

本文中出现的符号及其含义如表 1 所示：

表 1 符号说明

符号	含义
$H_0$	原假设
$H_1$	备选假设
$r$	每行所包含的参数数量
$c$	每列所包含的参数数量
$\chi^2$	实际值与理论值偏差的绝对大小
$P$	卡方分布的临界概率
$X_i$	SVM 训练样本集
$x_i$	待训练的特征向量
$y_i$	分类的类别，即特征值
$m$	权重
$b$	偏值
$\xi$	分类误差

## 五、模型的建立与求解

### 5.1 数据预处理

#### 5.1.1 季节、年龄段和过敏性疾病划分

对于数据中日期、年龄按季节、年龄段进行划分，详情见表 2。

表 2 季节与年龄段划分

名称	阶段类别	阶段划分	划分依据
季节	春季	2 月 4 日 - 5 月 4 日	2013-2017 年立春、立夏、立秋、立冬的阳历时间基本一致，故取
	夏季	5 月 5 日 - 8 月 6 日	
	秋季	8 月 7 日 - 11 月 6 日	2013 年的立春、立夏、立秋、立冬时间作为季节的划分时间节点
	冬季	11 月 7 日 - 2 月 3 日	
年龄	少年	0 - 14 岁	根据中华人民共和国政府对年龄的划分
	青年	15 - 34 岁	
	中年	35 - 55 岁	
	老年	56 - 100 岁	

研究所给数据科室类别并查阅相关资料后，本文按医学界常见过敏性疾病类型的划分标准分为皮肤过敏、呼吸道过敏和消化道过敏三大类，详情见表 3，除皮肤科、呼吸科以及消化科明显对应属于皮肤过敏、呼吸道过敏和消化道过敏外，其余科室由于所给数据样本较小并根据其与三大过敏性疾病之间的相关性进行归类划分，理由如下：

表 3 过敏性疾病划分及理由

过敏性疾病	科室	病例数	划分理由
皮肤过敏	皮肤科	857	具有皮肤过敏的临床表征
	耳专病门诊	2	患者直接前往耳专病门诊而不是耳鼻喉科，可排除呼吸道过敏的可能，眼部皮肤过敏会前往眼科就诊，故划归为皮肤过敏
	眼科	33	
	妇科	5	由于进行了过敏原检测并且未直接前往皮肤科就诊，有理由认为患者应为某些隐私部位皮肤过敏、不便于前往皮肤科就诊，故划归为皮肤过敏
	泌尿科	1	
	康复医学科	3	由康复医学科诊疗项目判断，且内分泌失调等问题也易在皮肤上表征，患者混淆的可能性很大，该患者大概率为皮肤过敏。灰甲趾由于其病因和临床表征，也被划归为皮肤过敏
	内分泌科	30	
	灰甲趾	3	

皮肤过敏	普外科	2	由于过敏会引起的皮肤浮肿等症状，同时有可能
	肿瘤科	1	引起关节等部位不适，所以前往普外、肿瘤、推
	推拿科	1	拿、骨科就诊并进行过敏检测的患者被划归为皮
	骨科	2	肤过敏
	心内科	8	由于皮肤过敏引发的的心脏类问题如心率不齐等使 得患者忽略过敏前往心内科就诊，故划归为皮肤 过敏
	肾内科	18	过敏性紫癜有概率引发紫癜性肾炎，所以前往肾 内科且进行过敏原检测的患者划归为皮肤过敏
	中医	32	中医、中医科患者多为调养皮肤，因为呼吸道、 消化道过敏症状较为明显，患者一般不会优先选
	中医科	14	择中医的缓慢调理、而是西医尽快缓解病症，故 划归为皮肤过敏
	呼吸科	627	具有呼吸道过敏的临床表征
	耳鼻喉科	173	呼吸道过敏的临床表现多为轻者反复咳嗽、打喷 嚏、流鼻涕、鼻塞等，重者可表现为胸闷、呼吸 不畅、发生哮喘。同时，仅是耳朵出现不适的患 者不会进行过敏原检测，故耳鼻喉科、口腔的患 者全部被划归为呼吸道过敏
呼吸道过敏	口腔科	1	
	神经科	5	由于呼吸道过敏时呼吸、吞咽困难和神经炎的临 床表征相似，所以前往神经科且进行过敏原检测 的患者划归为呼吸道过敏
	消化科	7	具有消化道过敏的临床表征
消化道过敏	普内科	49	当患者由于腹泻、呕吐等问题并未意识到为过敏 时，前往消化内科（即普内科）和血液科就诊十 分常见，故划归为消化道过敏
	血液科	3	

### 5.1.2 数据选择

数据所给对象取自 2013 年 1 月 5 日至 2017 年 12 月 29 日期间某院过敏信息统计表，所给 2284 例病例，去除性别未知一例，实际可用数据 2283 例。其中男性 1014 例，占总病例数的 44.42%，年龄在 1-94 岁，平均年龄 40.4；女性 1269 例，占总病例数的 55.58%，年龄在 1-98 岁，平均年龄 41.38。根据问题一要求，从 2283 例有效数据中去除难以划分到各类过敏性疾病中的综合性科室，如儿科、体检科、预防中心等综合性科室共 257 例，问题一将基于表 2 各科室所包含的病例共 2026 例进行计算分析，其中检测结果全为阴性的共计 895 人，检测结果为

阳性的共计 1131 人。剔除数据理由如下：

(1) 所给数据 V 列存在 4 组冗余数据，故这四组数据作为异常组在数据预处理中删除；

(2) 性别未知 1 项，对于过敏性疾病与性别的判定存在干扰，故在数据预处理中删除；

(3) 前往体检中心、体检科进行过敏原检测的 236 位患者，其是否过敏存在较大不确定性，故在数据预处理中删除；

(4) 过敏原检测中存在空白项的患者有 17 位，无法对其进行赋值分析，故作为干扰项在数据预处理中删除；

### 5.1.3 数据数字化

为了方便后续的数据处理和分析，本文将数据集进行了数字化处理，具体来说分为两个部分：①使用表 3 中对应值进行数字化，其中 1、3、5、7 列为按照上文所述的方法将数据集中的性别、季节、年龄段、过敏性疾病进行分类；2、4、6、8 列为对应的数字化值。②对于过敏原结果，我们通过查找相关检验标准及不同表示形式的过敏原检测结果表（详见附录 1），并向在血液科从事多年化验工作的主治医师咨询求证，得到以下数字化方案：当结果为 0、0(-)、0.00 时表示结果为阴性，均记为 0；1(±)的值表示为弱阳性，我们尝试使用 0.5、0.6、0.7、0.8、0.9 等数字化值进行 5.3 节中 SVM 模型的拟合（具体结果见附录 2），发现当 1(±)的数字化值为 0.8 时模型拟合效果最好，故选取 0.8 作为 1(±)的数字化值；2(+)表示轻度过敏，故记为 2；3(2+)表示中度过敏，故记为 3；4(3+)表示高度过敏，故记为 4；5(4+)表示很高度的过敏，故记为 5；6(5+)表示非常高程度的过敏，故记为 6；其他值保持不变。

表 4 性别、季节、年龄段、过敏性疾病数字化值对应表

因素	赋值
性别	女=0，男=1
年龄段	少年=1，青年=2，中年=3，老年=4
季节	春季=1，夏季=2，秋季=3，冬季=4
过敏性疾病	皮肤过敏=1，呼吸道过敏=2，消化道过敏=3
年份	2013=1, 2014=2, 2015=3, 2016=4, 2017=5

## 5.2 问题一的求解

据题意可知，要求分别分析过敏性疾病与时间、季节、性别、年龄因素之间的关系。基于对所给数据的季节、年龄等因素的阶段划分，一是采用 $\chi^2$ 独立性检验对过敏性疾病与时间、季节、性别、年龄之间进行分析， $P < 0.05$  认为差异具有统计学意义；二是通过分别绘制出过敏性疾病与时间、季节、性别、年龄之间的示意图，将过敏性疾病与时间、季节、性别、年龄之间的关系进行可视化处理，



将统计数据值与相关性图表结合起来对过敏性疾病与性别、年龄、季节、时间的关系做出解释。

## 5.2.1 数据分析

### 5.2.1.1 过敏性疾病与性别、年龄、季节、时间的关系

根据所给 2013 年 1 月 5 日-2017 年 12 月 29 日数据显示，有 2026 人前往医院检查是否患有过敏性疾病，根据过敏原检测数据及结果显示共有 1131 人过敏原检测结果呈阳性（确诊过敏），其中男性过敏者的比例为  $517/1131=60.05\%$ ，女性过敏者的比例为  $614/1165=52.70\%$ ，从比例初步可以看出男性患过敏性疾病的比例略高于女性。问题一对过敏原检测数据建立列联表，采用  $\chi^2$  检验来检

(1) 分析方法：

首先建立不同类别的过敏性疾病患者频数四格表，如下表所示：

表 5 过敏性疾病患者与各因素的频数表

因素	阴性 $B_0$	阳性 $B_1$	合计
$A_0$	$n_{00}$	$n_{01}$	$n_{0\cdot}$
$A_1$	$n_{10}$	$n_{11}$	$n_{1\cdot}$
合计	$n_{\cdot 0}$	$n_{\cdot 1}$	$n$

其中：

A——因素，即性别（男，女）；年龄（少年，青年，中年，老年）；季节（春季，夏季，秋季，冬季）；时间（2013-2017，共 5 个年份）。

B——是否患过敏性疾病， $B_0$  为阴性（无过敏）， $B_1$  为阳性（过敏）。

表 5 与表 6 中，当  $P_0=P_1$  时，A 与 B 相互独立，即各因素与过敏性疾病相互独立。以性别为例：

首先建立原假设  $H_0$ ：过敏性疾病与性别无关，即 A 和 B 是相互独立的。

故备选假设  $H_1$ ：过敏性疾病与性别有关，即 A 和 B 不相互独立。

由 Pearson 的  $\chi^2$  定理，有检验统计量

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(n_{ij} - n\hat{p}_{ij})^2}{n\hat{p}_{ij}} \quad (1)$$

在原假设  $H_0$  成立时，上式服从自由度为： $rc - (r+c-2) - 1 = (r-1)(c-1) = (2-1)(2-1) = 1$  的  $\chi^2$  分布。当  $P < 0.05$  是，说明过敏性疾病与性别之间的差异存在统计学意义，拒绝原假设  $H_0$  接受备选假设  $H_1$ 。

(2) 检验结果

将过敏性疾病病例人数分别与性别、年龄、季节、时间进行  $\chi^2$  独立性检验，结果如表 5 所示，P 值均小于 0.05，所以拒绝原假设  $H_0$ ，接受备选假设  $H_1$ ，这说明过敏性疾病与性别、年龄、季节、时间之间的差异存在统计学意义。

表 6 过敏性疾病与各因素之间的 $\chi^2$ 独立性检验结果

因素	病例总数	阴性		阳性		$\chi^2$ 检验	
		例数	%	例数	%	$\chi^2$ 值	P 值
性别						10.824	<0.05
男性	861	344	39.95%	517	60.05%		
女性	1165	551	47.30%	614	52.70%		
总数	2026	895	44.18%	1131	55.82%		
年龄						69.081	<0.05
少年	287	83	28.92%	204	71.08%		
青年	619	229	37.00%	390	63.00%		
中年	522	264	50.57%	258	49.43%		
老年	598	319	53.34%	279	46.66%		
总计	2026	895	44.18%	1131	55.82%		
季节						22.008	<0.05
春季	452	221	48.89%	231	51.11%		
夏季	491	186	37.88%	305	62.12%		
秋季	558	226	40.50%	332	59.50%		
冬季	525	262	49.90%	263	50.10%		
总计	2026	895	44.18%	1131	55.82%		
时间						114.6	<0.05
2013	381	161	42.26%	220	57.74%		
2014	428	236	55.14%	192	44.86%		
2015	536	160	29.85%	376	70.15%		
2016	379	229	60.42%	150	39.58%		
2017	302	109	36.09%	193	63.91%		
总计	2026	895	44.18%	1131	55.82%		

### 5.2.1.2 三大过敏性疾病与性别和年龄的关系

运用对应分析的基本思想，分析二维数据矩阵中行因素和列因素间的关系。对应分析的基本原理是，对二维数据矩阵进行适当的变换（即对应变换），使变换后的数据对行与列是相对应的，从而可以同时行和列进行分析，以便发现行列因素间的关系。建立过敏性疾病与年龄之间的列联表，研究过敏性疾病与年龄之间的关系。

#### （1）分析与讨论

首先建立过敏性疾病人数和年龄之间的二维频数表，见表 7。将表 7 中的 3 种过敏性疾病作为样本（ $Y_j, j=1,2,3$ ），4 个年龄段作为变量（ $X_i=1,2,3,4$ ）对

2013年1月5日~2017年12月29日该医院过敏性检测情况进行对应分析，旨在探讨这些过敏性疾病与年龄之间的关系。

表 7 三种过敏性疾病各年龄段过敏例数

年龄段	皮肤过敏 $Y_1$	呼吸道过敏 $Y_2$	消化道过敏 $Y_3$	合计
少年(1)	67	26	0	93
青年(2)	179	194	17	390
中年(3)	117	125	16	258
老年(4)	147	129	3	279
合计	510	474	36	1020

## (2) 分析结果

将表7转化为行变量百分比分布表，它将表7中的原始频数分布转换为基于行变量的百分比分布，如下表8所示，其中Mass是一个基于边际频数的影响量，可称为众位数，为列或行百分比的加权均数。Mass越大，对形心

(centroid)的影响越大，也就是越靠近形心。表8中，从列的Mass可以看出，皮肤过敏与呼吸道过敏较靠近形心；从行的Mass可以看出，青年的Mass最大，其次为中年和老年。而少年由于本次消化道过敏人数为0，因此不具有参考价值。

表 8 三种过敏性疾病与年龄段之间的行列变量百分比分布表

年龄段	皮肤过敏	呼吸道过敏	消化道过敏	Mass
少年	0.07	0.03	0.00	0.091
青年	0.18	0.19	0.02	0.382
中年	0.11	0.12	0.02	0.253
老年	0.14	0.13	0.00	0.274
Mass	0.50	0.47	0.04	1.000

图1为三大过敏性疾病与年龄的对应分析因子负荷图，图2为过敏性疾病与年龄的相关性分析图，将两张图结合进行分析：由图1可以看出，呼吸道过敏紧靠青年，其次为中年和老年，说明呼吸道过敏与青年有很强的关系，即青年人更易得呼吸道过敏性疾病；

皮肤过敏与老年的距离最近，说明过敏性疾病与年龄有一定的联系。图2中可以看出，不同的年龄段中，青年人患过敏性疾病的人数占比重最大，验证了图1中结果青年人更容易患过敏性疾病。

分析其原因，可能是由于青年人长期处于密闭的写字楼等环境中，加上使用中央空调。中央空调如果消毒不到位的话，引起过敏的物质很有可能随着空调四处循环，加上青年白领室外活动的机会减少，从而导致新鲜空气呼吸范围减少。并且可以看出女性患过敏性疾病的概率基本处于男性水平线上，这可能是由于女性的皮肤较男性更为脆弱敏感而更易引起过敏导致。女性工作者的工作性质更偏向于室内，所以女性工作者相较男性工作者更爱长期在办公室办公而不外出办公，所处环境的空气不流通等原因导致<sup>[3]</sup>。

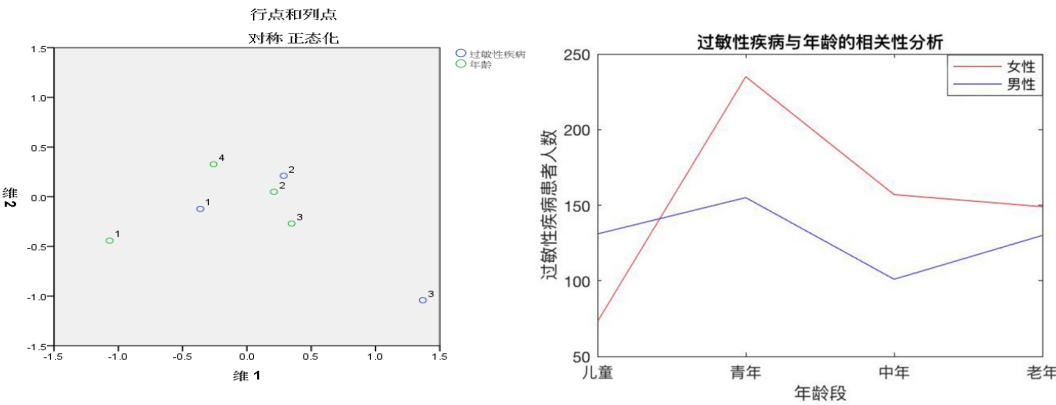


图 1 三大过敏性疾病与年龄的对应分析因子负荷图      图 2 过敏性疾病与年龄的相关性分析图

### 5.2.1.3 三大过敏性疾病与性别和年龄的季节的关系

#### (1) 分析与讨论

首先建立过敏性疾病人数和季节之间的二维频数表，如表 9 所示。将表 9 中的 3 中过敏性疾病作为样本 ( $Y_j, j=1,2,3$ )，4 个季节段作为变量

( $X_i=1,2,3,4$ ) 对 2013 年 1 月 5 日~2017 年 12 月 29 日该医院过敏性检测情况进行对应分析，旨在探讨这些过敏性疾病与季节之间的关系。

表 9 三种过敏性疾病各季节过敏例数

季节	皮肤过敏 $Y_1$	呼吸道过敏 $Y_2$	消化道过敏 $Y_3$	合计
春季(1)	111	93	8	212
夏季(2)	138	126	8	272
秋季(3)	153	130	10	293
冬季(4)	108	125	10	243
合计	510	474	36	1020

将表 9 转化为行变量百分比分布表。它将表 9 中的原始频数分布转换为基于行变量的百分比分布，如下表 10 所示。在这里，从列的 Mass 可以看出，皮肤过敏与呼吸道过敏较靠近形心；从行的 Mass 可以看出，Mass 在秋季最大<sup>[4]</sup>。

表 10 三种过敏性疾病与季节之间的行列变量百分比分布表

季节	皮肤过敏	皮肤过敏	皮肤过敏	Mass
春季	0.11	0.09	0.01	0.208
夏季	0.14	0.12	0.01	0.267
秋季	0.15	0.13	0.01	0.287
冬季	0.11	0.12	0.01	0.238
Mass	0.50	0.47	0.04	1

## (2) 分析结果

图3为三大过敏性疾病与季节的对应分析因子负荷图，图4为过敏性疾病与季节的相关性分析图，将两张图结合进行分析：由图3可以看出，皮肤过敏与秋季的距离最近，说明皮肤过敏在秋天发生的可能性较大。图4中可以看出，不同的季节中，秋季患过敏性疾病的人数占比最大，验证了图3中结果秋季更容易患过敏性疾病。

分析其原因，可能是由于秋天空气干燥，气候变化频繁，紫外线增强等原因导致秋天过敏性多发，尤其是皮肤过敏。这个季节皮肤过敏最常出现的就是过敏性皮炎、颜面再发皮炎等，这类患者尤以女性居多，从图4可以看出女性秋天皮肤过敏的概率大于男性，常见的反应如颜面、颈部等暴露部位出现红斑、水肿、丘疹、细薄脱屑，或者表面轻度增厚，有的表现为眼周或颈部水肿性红斑，常伴有瘙痒或刺痛。

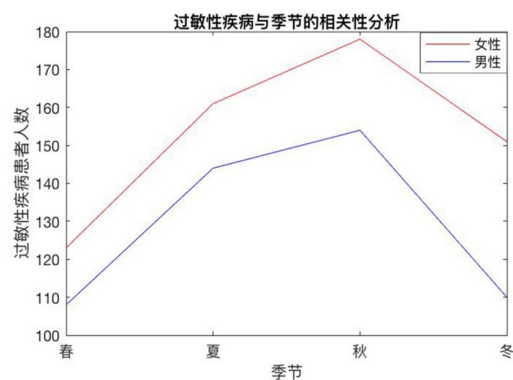
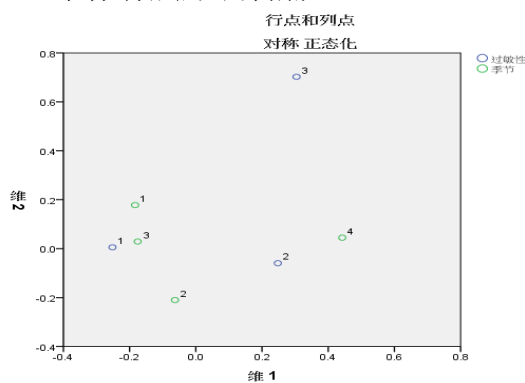


图 3 三大过敏性疾病与季节的对应分析因子负荷图

图 4 过敏性疾病与季节的相关性分析图

## 5.2.1.4 三大过敏性疾病与性别和时间的关系

### (1) 分析与讨论

首先建立过敏性疾病人数和时间之间的二维频数表，如表 11 所示。将表 11 中的 3 中过敏性疾病作为样本 ( $Y_j, j=1,2,3$ )，5 个年份时间段作为变量 ( $X_i=1,2,3,4,5$ ) 对 2013 年 1 月 5 日~2017 年 12 月 29 日该医院过敏性检测情况分为 5 各年份进行对应分析，旨在探讨这些过敏性疾病与时间之间的关系。

表 11 三种过敏性疾病与年份之间的过敏例数

时间	皮肤过敏 Y1	呼吸道过敏 Y2	消化道过敏 Y3	合计
2013(1)	156	39	8	203
2014(2)	75	91	5	171
2015(3)	147	178	10	335
2016(4)	48	83	3	134
2017(5)	84	83	10	177
合计	510	474	36	1020

将表 11 转化为行变量百分比分布表。它将表 11 中的原始频数分布转换为基于行变量的百分比分布，如下表 12 所示。在这里，从列的 Mass 可以看出，皮肤过敏与呼吸道过敏较靠近形心；从行的 Mass 可以看出，Mass 在 2015 最大。

表 12 三种过敏性疾病与时间之间的行列变量百分比分布表

时间	皮肤过敏	呼吸道过敏	消化道过敏	Mass
2013	0.15	0.04	0.01	0.199
2014	0.07	0.09	0.00	0.168
2015	0.14	0.17	0.01	0.328
2016	0.05	0.08	0.00	0.131
2017	0.08	0.08	0.01	0.174
Mass	0.50	0.47	0.04	1

## (2) 分析结果

由图5可以看出，消化道过敏在2014、2015、2016年间过敏病例数较高，原因可能是由于2015年由于该地生态环境遭到破坏、大气污染，导致许多过敏原的数量达到峰值，例如树组合、普通豚草、艾蒿、屋尘、猫毛等，并且从图6可以看出女性过敏病例基本处于男性之上，结合之前的分析结果证明过敏性疾病与性别之间存在一定的关系。

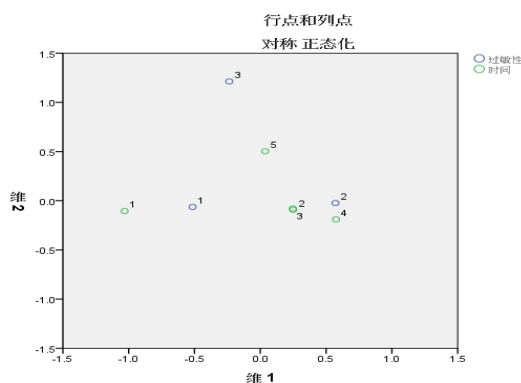


图 5 过敏性疾病与时间的对应分析因子负荷图

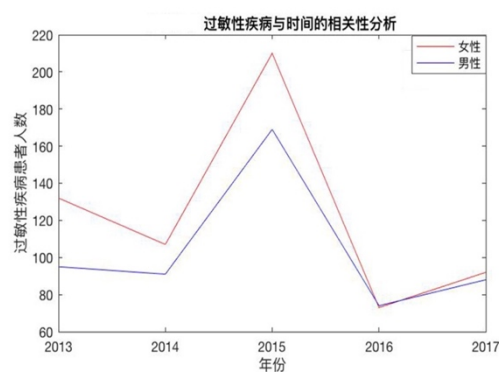


图 6 过敏性疾病与季节的相关性分析图

根据上述分析结果，将过敏性疾病与性别、年龄、季节、时间之间存在的  
关系及缘由总结如下表 13 所示：

表 13 过敏性疾病与性别、年龄、季节、时间的关系总结

因素	与过敏性 疾病独立 否	关系	可能缘由
性别	否	女性过敏的可能 性高于男性，尤 其是皮肤过敏	女性的皮肤较男性更为脆弱敏感而 更易引起过敏导致
男			
女			
年龄	否	青年人更容易患 过敏性疾病，尤 其是呼吸道过敏 性疾病	青年人长期处于密闭的写字楼等环 境中，长期使用的中央空调如果消毒 不到位的话，引起过敏的物质很有可 能随着空调四处循环，加上青年白领 室外活动的机会减少，尤其是女性， 从而导致新鲜空气呼吸范围减少
少年			
青年			
中年			
老年			
季节	否	秋天患过敏性疾 病的 可 能 性 较 大，尤其是皮肤 过敏	秋天空气干燥，气候变化频繁，紫外 线增强等原因导致秋天过敏性多发， 尤其是皮肤过敏，且以女性患者居多
春季			
夏季			
秋季			
冬季			
时间	否	2015 年较多人患 过敏性疾病	根据过敏原检测结果猜测可能是由 于 2015 年由于该地生态环境遭到破 坏、大气污染，导致许多过敏原的数 量达到峰值，例如树组合、普通豚草、 艾蒿、屋尘、猫毛等
2013			
2014			
2015			
2016			
2017			

### 5.3 问题二的求解

根据题意，问题二要求结合第一问的相关影响因素，分析不同类型疾病的病人过敏原检测结果之间的差异。为了更好的寻找题设所求的差异，我们引入了机器学习中的 SVM（Support Vector Machine,支持向量机）模型，这是最近火热的机器学习中常用的分类模型，该算法基于统计学习理论基础，可以有效地对每个数据进行分类，并寻找到用于分类的超平面，从而对其他过敏性疾病种类未知的患者进行预测。另外，该模型可以根据对数据的分类输出预测变量的重要性，我们可以通过对该重要性进行分析，进而得到题设所求的差异。然后，我们结合第

一间的分析结果与上述模型得到的预测变量的重要性,使用贝叶斯网络模型对主要特征进行分析,从而得到过敏原与季节、性别、年龄之间可能的联系。

### 5.3.1 SVM 模型说明

SVM 是一种泛化能力很强的分类模型,其在解决小样本分类问题方面效果明显,其基本思想是在样本集线性不可分的情况下,将输入的特征向量通过非线性变换映射到另一个更高维度的空间,从而在变换后的高维空间中找到一个最优的超平面进行分类,这种方法可以在将两类数据分类的同时使得两类的分类间隔最大。现对训练样本集  $X_i = (x_i, y_i)$  使用 SVM 算法进行分类,其中  $x_i$  表示待训练的特征向量,  $y_i$  表示分类的类别,即特征值。首先对训练样本集  $X_i$  做非线性变换:  $Z = \phi(X_i)$ ,使得训练样本集  $X_i$  变得线性可分,即可以通过训练找到权重  $m$  和偏值  $b$ ,满足  $m^T Z + b = 1$ , 或者  $m^T Z + b = -1$ ,再假设最优的超平面方程  $H$  为:

$$m_0^T Z + b_0 = 0 \quad (2)$$

对于已经分好类的数据,特征向量与  $H$  之间的距离可以表示为:

$$p(m, b) = \min_{\{x|y=1\}} \frac{z^T m}{\|m\|} - \max_{\{x|y=-1\}} \frac{z^T m}{\|m\|} \quad (3)$$

因此,当

$$p(m_0, b_0) = 2/\|m_0\| = 2/\sqrt{m_0^T m_0} \quad (4)$$

最大时,求最优超平面方程  $H$  的问题可以被转化为一个常规的求解二次规划的问题:

$$\min_{w,b} \phi(m) = \frac{1}{2} (m^T m) \quad (5)$$

对于线性不可以分的情况,需求出:

$$\min_{w,b} \phi(m) = \frac{1}{2} m^T m + \gamma \sum_{i=1}^n \xi_i \quad (6)$$

的最小值,其中  $\xi$  为分类的误差。对于上式,利用拉格朗日乘子法求解,可得:

$$m_0 = \sum_{i=1}^l \lambda_i y_i Z_i \quad (7)$$

进而可以得出分类函数:

$$f = \text{sgn} \sum_{i=1}^{20} \lambda_i y_i (Z^T Z_i) + b \quad (8)$$

在本题中,针对过敏性疾病的分类问题是一个多类的问题。因此,需要对上述传统的 SVM 算法进行改造,构造 SVM 多类分类器主要有两种方法,分别是 OVR(one-versus-rest) SVMs 法和 OVO(one-versus-one) SVMs 法,本文选用的是 OVO(one-versus-one) SVMs 法,即在任意两个样本之间设计一个 SVM 分类器,因此若需对  $K$  个类别的数据进行分类就需要训练  $K(K-1)/2$  个 SVM 分类器,当需要对一个数据进行分类时,每个 SVM 分类器都对其进行类别的判断,并对判断出来的类别进行投票,最后汇总每个 SVM 分类器投票的类别,得到票数最高的类别即为该数据的类别。该方法相对于 OVR SVMs 的优势在于运行测试模型



时能较快的得出结果。在本文中，共有 3 个类别的数据，则使用 OVO SVMs 法需要训练 3 个 SVM 分类器。

5.3.2SVM 模型构建

选用 SPSS Modeler 14.1 用于构建 SVM 模型，该软件可以通过一个 SVM 模块对数据使用 OVO SVMs 法进行分类，这样做不需要多次调试 SVM 分类器，从而加快模型整体的建立速度。在本模型中，训练数据中仅使用前文数据中的皮肤、呼吸道、消化道三类，并把训练数据分为“训练”和“测试”两个分区，所占比例分别为 70% 和 30%，所使用的数据分别为 785 条和 346 条。在设置中选择随机种子数为 815318，使得上述分区中的数据是随机分配的。选择专家建模选项，经过多次实验，选取模型预测结果最高的参数，其中设置的停止标准为 $1.0\times e^{-3}$ ，规则化参数为 10，回归精确度为 0.1，内核类型为多项式。

5.3.3SVM 模型预测效果及分析

选用 Modeler 的分析节点对模型的整体性能进行评价，从而输出模型的预测结果矩阵，如表 1 所示。分析该表可知，本文训练的 SVM 模型总体预测正确率为 93.35%。接着本文使用该模型对儿科、体检科和体检中心等科室的患者进行了预测，具体见附录 3，附录 4，其中过敏类型列表示模型推断出患者可能患的疾病，该结果可以为患者选择下一步诊疗方案提供参考。我们建立上述模型并进行预测，希望可以通过该方式为获得过敏原报告的患者提供可能患有的疾病的预警，从而提出建议让患者提前防范该类疾病，以减少过敏性疾病带给患者的痛苦。

表 14 模型的预测结果

分区	训练集		测试集	
正确	737	93.89%	323	93.35%
错误	48	6.11%	23	6.65%
总计	785	100%	346	100%

该模型在进行分类后输出了一个重要的指标，即模型中预测变量的重要性，如图 7 所示，即在所有预测变量中，对输出变量分组影响从小到大的排序，其中前六个变量为：律草、年龄段、蟹、蟑螂、普通豚草和霉菌组合。即对于不同类型疾病的病人过敏原检测结果差异最大的前五个过敏原别为：律草、蟹、蟑螂、普通豚草和霉菌组合。

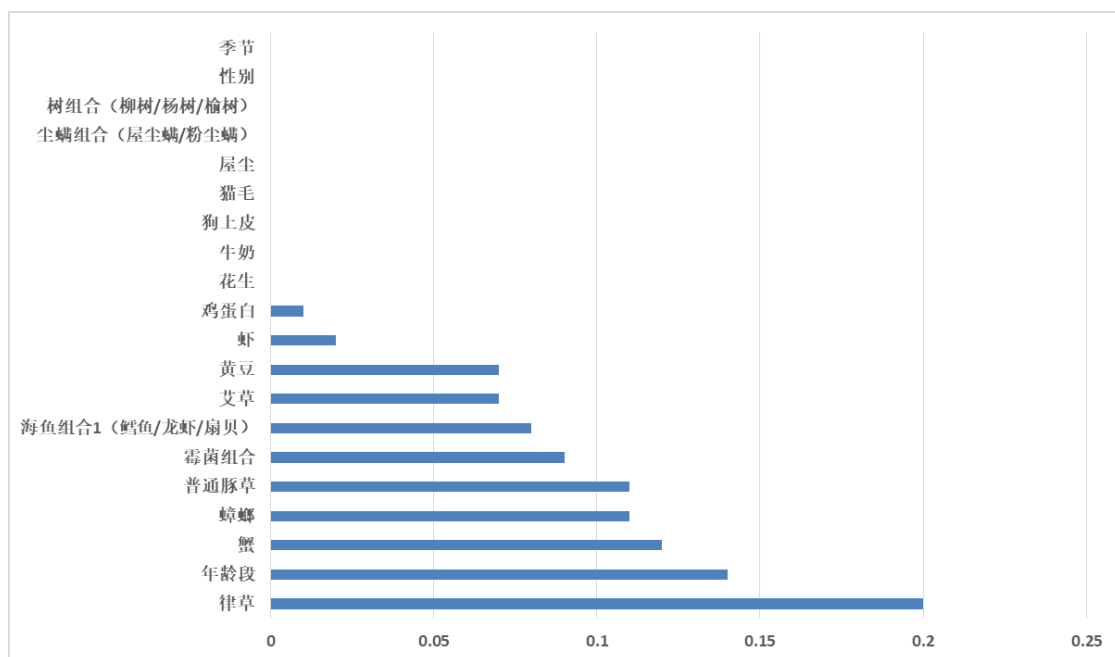


图 7 预测变量的重要性

### 5.3.4 贝叶斯网络模型说明

为了分析数据集中主要变量之间的关系，我们选用贝叶斯网络模型进行分析。该模型是数据分析中的一种算法模型，主要用于对数据之间的关系进行分析和处理。该网络模型是以贝叶斯方法作为基础，使用贝叶斯概率进行运算得到的。由于数据集中输入变量太多，为了节省模型运算时间、便于分析，本文对数据进行特征预处理。在上述 SVM 模型，得到的五大过敏原为：律草、蟹、蟑螂、普通豚草和霉菌组合，在第一问中，我们发现季节、性别、年龄与过敏性疾病均有较大的相关性，故选用律草、蟹、蟑螂、普通豚草、霉菌组合、季节、性别、年龄等八个特征作为输入。

### 5.3.5 贝叶斯网络模型构建

选择 SPSS Modeler 14.1 用于构建贝叶斯模型，选取上述八个特征与过敏性疾病进行分析，模型选择的是专家建模选项，使用似然比进行独立测试。这里使用的训练数据中依旧使用前文数据中的皮肤、呼吸道、消化道三类，但因只需要获取变量之间的相关性，故使用全部数据集进行测试，而不像 SVM 模型那样对数据进行分区，最后输出的贝叶斯网络模型见图 8。

## 贝叶斯网络

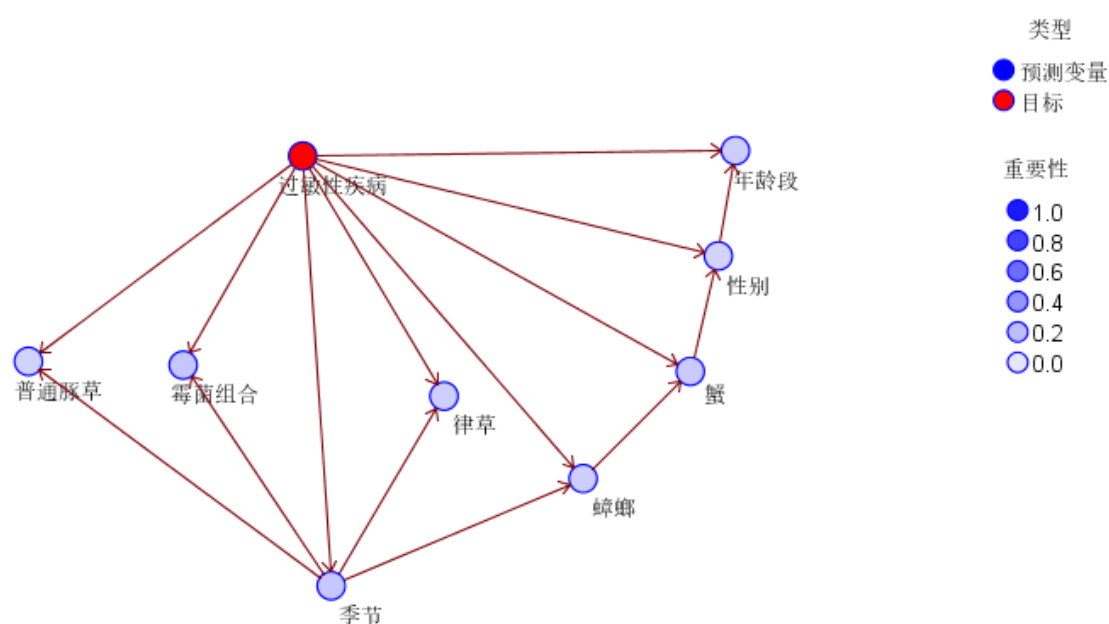


图 8 贝叶斯网络模型结构图

### 5.3.6 贝叶斯网络模型效果及分析

根据构建的模型图可以看到，季节是普通豚草、霉菌组合、律草和蟑螂的父节点，说明季节对过敏性疾病的影响不仅取决于自身，还依赖于这三个变量；蟹是性别的父节点，说明蟹对过敏性疾病的影响不仅取决于自身，还依赖于这性别这个变量。从上所述可知①普通豚草、霉菌组合、律草和蟑螂等过敏原对过敏性疾病的影响与季节关系较大。该分析与事实相吻合，每年夏季是豚草和律草过敏的高发期；而霉菌组合和蟑螂因为环境因素在冬季时数量较小，继而影响过敏性疾病发生的数量②蟹这个过敏原对过敏性疾病的影响与性别有关。该分析与事实较吻合，使用百度搜索“男 蟹过敏”共有 2,010,000 条结果，而搜索“女 蟹过敏”则高达 3,350,000 条，数量是男性的约 1.7 倍，远超过搜索“女 过敏”（6,710,000 条）和搜索“男 过敏”（5,830,000 条）的结果比例（1.1 倍），说明性别对蟹这个过敏原所带来的疾病影响较大。

### 5.4 问题三的求解

据题意可知，要去定量分析从 2013 年到 2017 年过敏原检测结果的变化趋势，选取一个环境保护指标并分析两者之间的相关性。首先定量分析 2013-2017 年内 17 种过敏原阳性率，作出其变化趋势示意图并分析其变化趋势，然后对选取的环境保护指标在 2013-2017 年的变化趋势进行分析，最后两者的相关性进行评价。

根据题目提供数据，我们定量分析了 17 种过敏原在 2013-2017 年检测结果，具体如下图所示：

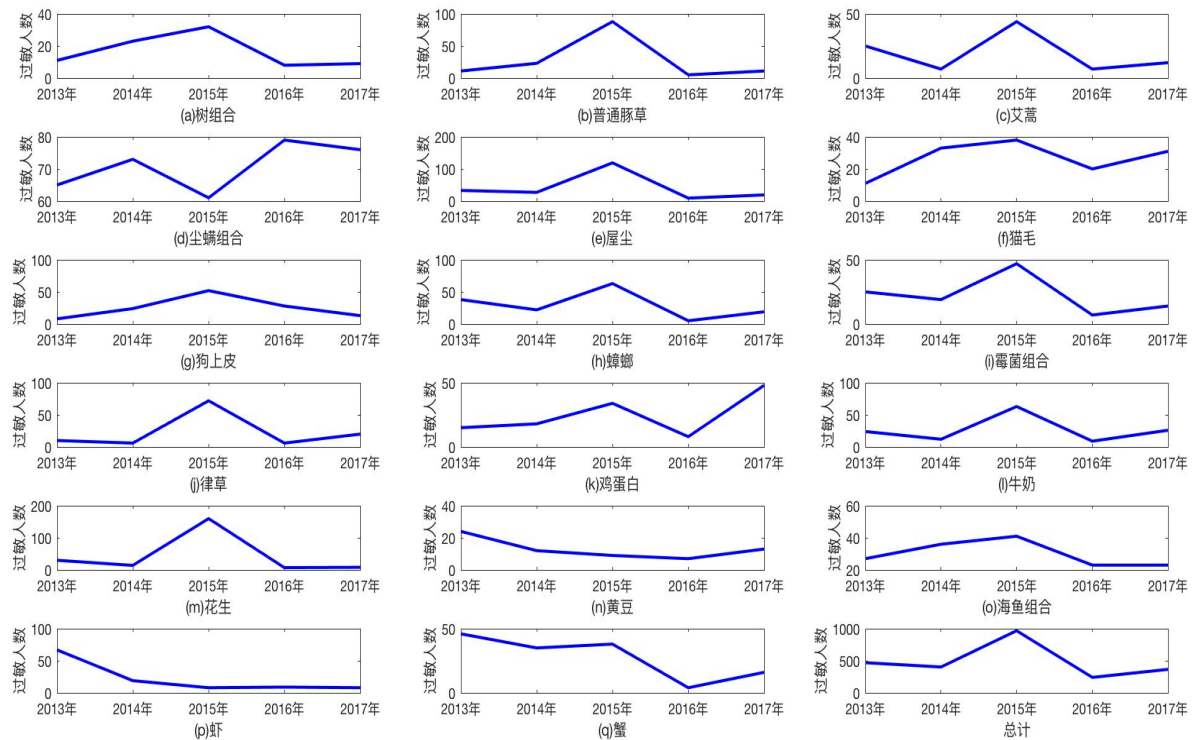


图 9 2013-2017 年过敏原检测结果变化趋势

由图 9 可直观看出 2013-2017 年间 17 种过敏原检测结果的变化趋势。其中，(a)树组合、(b)普通豚草、(c)艾蒿、(j)律草这四种吸入性过敏原的检测结果有着非常相近的变化趋势。考虑到树组合、豚草、艾蒿、律草都为植物类过敏原，且其诱发过敏的方式大多都通过患者吸入花粉从而诱发呼吸道过敏，所以本题针对树组合、普通豚草、艾蒿、律草这四种过敏原诱发呼吸道过敏进行分析。

CO<sub>2</sub>作为最主要的温室气体，在环境保护指标中占有较为重要的地位。大气中的CO<sub>2</sub>含量由于植物生长而产生随季节变化的特性。由于CO<sub>2</sub>具有较好的保温性，故可引起全球气温升高。

根据中华人民共和国生态环境部提供的环境保护指标，结合本题选择对植物类吸入性过敏原的变化趋势进行分析，故选取温室气体中的CO<sub>2</sub>浓度作为环境保护指标，与植物类吸入性过敏原进行相关性分析。

根据中国气象局 2013-2017 年发布的《中国温室气体公报》数据显示，2013-2017 年我国陆地区域年平均大气CO<sub>2</sub>浓度如下表所示：

表 15 2013-2017 年中国陆地区域年平均 CO<sub>2</sub> 浓度

年份	2013	2014	2015	2016	2017
浓度/ppm	397.3	398.7	401	404.4	405.8

注：ppm 是用溶质质量占全部溶液质量的百万分比来表示的浓度，也称百万分比浓度。

为直观分析CO<sub>2</sub>与四种植物类吸入性过敏原的相关性，将其 2013-2017 年变化趋势折线图整合如下：

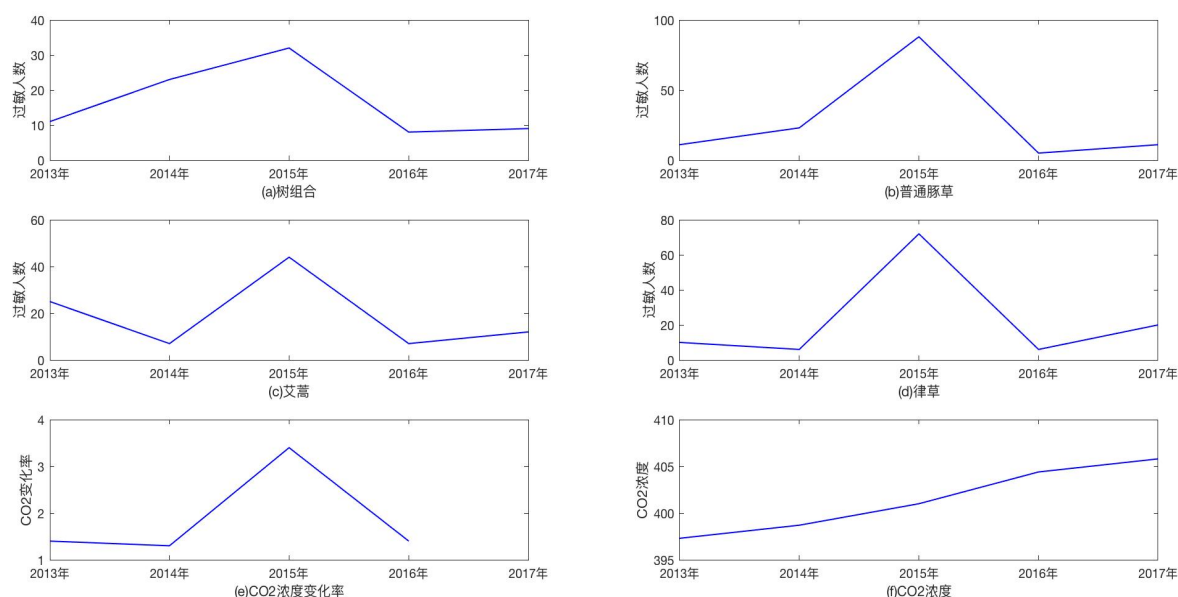


图 10 2013-2017 年CO<sub>2</sub>与过敏原的相关性分析

由于中国陆地区域年平均CO<sub>2</sub>浓度呈逐年上升趋势，故在(e)中作出其浓度变化率随时间的变化趋势。因为 2018 年的《中国温室气体公报》尚未发布，故 2017 年CO<sub>2</sub>浓度变化率数据暂时不明。

对比图 10 中的(a)-(e)，不难发现其变化趋势非常相近。为分析过敏原检测结果变化与CO<sub>2</sub>的相关性，本题主要讨论 2015 年前后的植物类吸入性过敏原检测结果与CO<sub>2</sub>浓度的关系。由于吸入性过敏原诱发的为呼吸道过敏，可认为诱发呼吸道过敏的案例数可反映出植物的数量、花粉量、花粉的传播效果这三大因素。2014 年间，由于CO<sub>2</sub>浓度增长较缓，故植物类过敏原中，只有数量较为庞大的树组合诱发过敏的数据较大，豚草、艾蒿、律草诱发的过敏案例较少。在 2015 年时，CO<sub>2</sub>浓度变化率明显高于其他年份，即该年CO<sub>2</sub>浓度的增长显著高于其他年份。与此同时，树组合、豚草、艾蒿、律草诱发的过敏人数也在 2015 年达到最高点。豚草、艾蒿、律草诱发的过敏性疾病案例显著提高，可反推得豚草等入侵性植物数量、花粉量增大、花粉传播效果更好。同时，树组合诱发过敏数也达到最高。然而随着入侵性植物数量的增加，树组合数量应减少或大致维持不变。因此可得出 2015 年植物花粉产量增长、花粉传播效果提高。

根据郭磊<sup>[5]</sup>等人的研究表明，温度升高可促使花药脱水和散粉加快，即温度升高对于植物花粉量的增加有着促进作用。武喆<sup>[6]</sup>等人研究发现，温度的升高对于花粉的扩散有着显著的促进作用。Christine Rogers<sup>[7]</sup>等人研究发现，不同浓度的CO<sub>2</sub>对植物花粉产生量的影响具有正相关性，随着温室气体的增加，人类将

迎来植物花粉暴露量增加 202% 的情形。由于植物大多于春、秋季开花、传播花粉，故温度难以达到抑制花药散粉、花粉活性的阈值。所以根据相关资料和本题数据分析，可得出 2015 年空气中  $\text{CO}_2$  浓度应有显著的增加，才会导致温度升高、植物光合作用增强。

根据图 10 中(e)、(f)的数据，可验证本题的推测符合数据变化趋势，具有一定的合理性。2016 年作为“十三五规划”的开局之年，“创新、协调、绿色、开放、共享”五大发展理念深入人心，各地牢记“绿水青山就是金山银山”这一绿色发展理念开展环保工作并初见成效。由于政策的影响，2016 年植物类吸入性过敏原检测阳性率明显降低，全年  $\text{CO}_2$  浓度增长率也降至低点。2017 年随着  $\text{CO}_2$  浓度持续升高，树组合、豚草、艾蒿、律草的过敏原检测阳性率不可避免地升高，但较 2015 年相比缓慢。

综上，本题选取的环保指标—— $\text{CO}_2$  浓度与过敏原检测结果具有正相关。

## 5.5 问题四的求解

根据题意可知，第四问需要结合本题数据、分析结果，针对可能诱发过敏性疾病的环保指标、我国政策国情进行生态环境保护倡议。针对题中关于环保指标的倡议，我们又分析了不同过敏原在不同季节的诱发过敏情况，以便更好地向市民提出针对、有效的倡议。

我们通过分别对过敏原和其在不同季节诱发过敏性疾病的情况进行分析，过敏原在不同季节的诱发过敏情况如图 11 所示。

可见树组合、豚草、艾蒿过敏在春夏两季诱发过敏情况较为严重；尘螨组合、屋尘、蟑螂、霉菌组合导致过敏在秋季集中爆发；猫毛、狗上皮在春秋猫、狗换毛时频发；律草引起的过敏主要集中在冬季到春季；鸡蛋白、牛奶、花生、黄豆、海鱼组合、虾、蟹等食入性过敏原与生态环保指标的相关性较弱，故不予讨论。



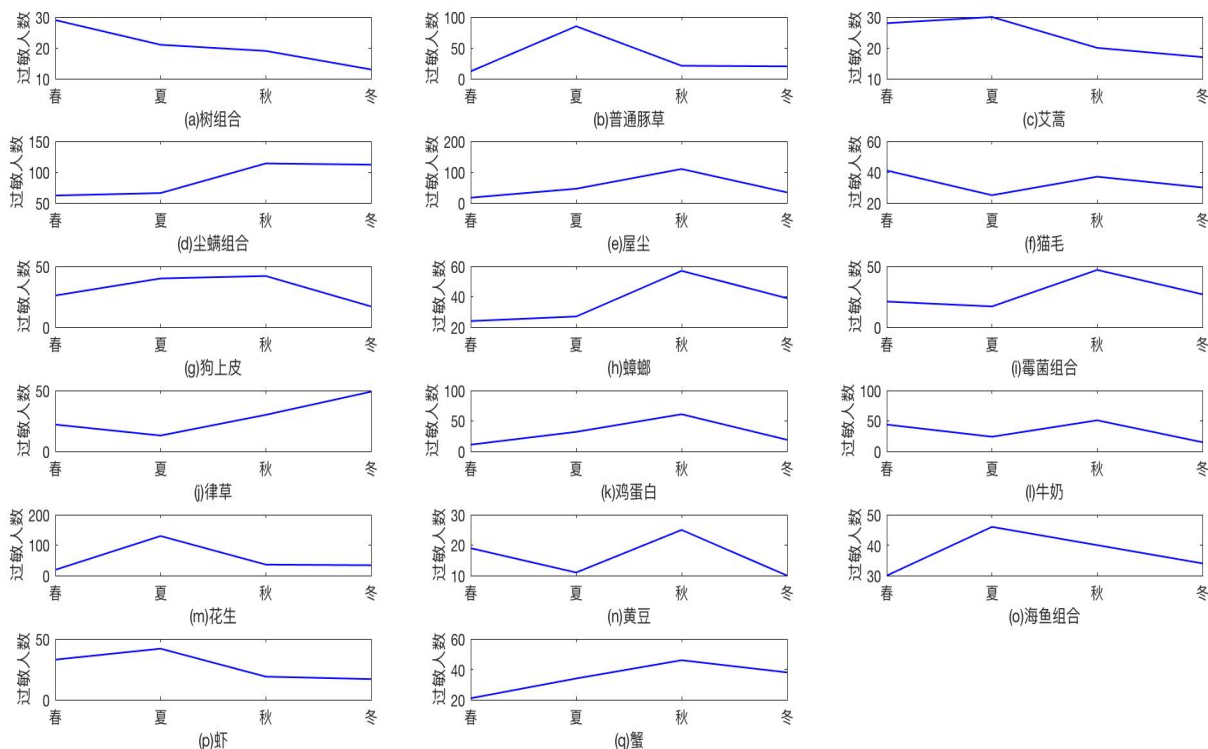


图 11 不同季节过敏原检测结果变化趋势

结合本文中的数据、分析，倡议书如下：

亲爱的市民朋友们：

大家好！

还没享受够满园飘香的盎然春意，转眼就感觉到了毒辣的太阳和火热的温度，夏天，已经大张旗鼓地来了。

每到这个时候开始，夜晚的街边总能看到成群结队的市民朋友吃龙虾、烧烤，享受着大自然的馈赠。龙虾、烧烤固然美味，但市民朋友们在品尝美食时，也要当心食物过敏。夏季为虾、海鱼、花生等食品过敏的高发季，这和饮食习惯有着很大的关系，喝着小酒，吃着海鲜烧烤小龙虾，无比惬意的同时，也给环境带来了极大的负担。传统的明火烧烤固然好吃，但也会产生大量的  $\text{CO}$ 、 $\text{CO}_2$ 、 $\text{SO}_2$  等气体，不仅污染大气，还会带来一系列的连锁反应。

近年来，呼吸道过敏已经屡见不鲜。暴露在空气中的花粉、动物皮毛组织等过敏原使得过敏患者苦不堪言，每天只能“全副武装”才敢出门。大家只知道罪魁祸首是那些花草树木，却不知道  $\text{CO}_2$  才是真正的幕后黑手吧？科研实验发现， $\text{CO}_2$  浓度对花粉产量、传播效率都有着很好的促进作用。同时， $\text{CO}_2$  又在导致全球变暖的温室气体中占有重要地位，可见其罪大恶极。

为了改善大气条件、缓解全球变暖趋势、减小呼吸道过敏疾病的发病概率，在此我发出倡议：

1.提倡无烟烧烤。市面上很多无烟烧烤设备，市民朋友在选择烧烤摊、自主烧烤时，尽可能地选择低污染、环保的无烟烧烤设施。在享受美食的同时也为生态环境保护作出了自己的努力；

2.摊主朋友们更新设备时选择无烟、环保型设备。不仅可以当作宣传噱头，还积极响应习总书记“绿水青山就是金山银山”的著名论断，实现摊位的“可持续发展”；

3.烧烤所用签尽可能使用金属制签替换木制签。不仅可以降低成本，还可以减少树木的砍伐，为保障植被的覆盖率、抑制温室气体的疯涨贡献自己的力量；

4.出行尽可能选择公共交通工具。不仅缓解交通拥堵、避免了道路问题导致的心情烦躁，还减少了碳排放，真正做到“保护环境，人人有责”。

5.坚持每天公益行走。不仅可以锻炼身体，还可以将步数进行公益捐赠、支付宝积攒能量种树等，通过“互联网+”公益的方式支持国家的环保政策。

习总书记一直强调“绿水青山就是金山银山”。我们作为努力实现中国梦的一分子，只要平时注意生活细节、养成良好的生活习惯，以自己的实际行动减少温室气体排放，我们相信，在不久的将来，我们不仅有希望摆脱花粉过敏等过敏性疾病的困扰，中国也将成为“富强、民主、文明、和谐、美丽”的社会主义现代化强国。

## 六、模型的检验与改进

### 6.1 考虑算法的拟合和泛化特性

在本文 5.1.3 节数据数字化中， $1(\pm)$ 的数字化值假设为 0.8，但实际上  $1(\pm)$ 并没有具体的量化指标，故我们的假设可能会影响模型的泛化效果。

本文假设各个过敏原是相互独立的，但是事实上有些过敏性疾病可能是多个过敏原与病毒等因素共同叠加所产生的结果，这可能会对模型的泛化效果产生影响。需要使用其他方法对其共同作用的可能性进行分析。

在 SPSS Modeler 软件中，只能实现含有一个隐藏层的神经网络，而本题所设的输入特征值较多，单隐藏层神经网络拟合效果较差（详见表 15），根据以往的经验，多层神经网络对数据的分类及泛化能力将会好明显好于单隐藏层神经网络。

### 6.2 更加多样化和详尽的数据

按照欧蒙医学实验诊断股份公司吸入性及食物性过敏原特异性 IgE 抗体检测试剂盒说明书说明，过敏包含如下几种：ts20(树组合),w1(普通豚草),w6(艾



蒿),ds1(尘螨组合),hl(屋尘),el(猫毛),e2(狗上皮),i6(蟑螂),msl(霉菌组合),u80(律草),f1(鸡蛋白),22(牛奶),f13(花生),f14(黄豆),f27(牛肉),f88(羊肉),fs33 (海鱼组合),f34(淡水鱼组合),f24(虾),123(蟹)。而本题提供的数据缺失了其中部分过敏原,缺失的过敏原可能与过敏性疾病有关,可能影响分析结果。

本题所给的数据中只提供了患者选择科室的数据,若提供患者所得的具体疾病类型,可以更加细化的预测患者所得的具体疾病,从而得到更好的结果。

表 16 单隐藏层神经网络拟合效果

分区	训练集		测试集	
正确	439	56.42%	184	53.38%
错误	346	43.58%	162	46.62%
总计	785	100%	346	100%

## 七、模型的评价与推广

### 7.1 模型的优点

- 1、问题一从“总—分”的角度分析过敏性疾病与各因素间的关系,并且图表与文字结合更加清晰易懂;
- 2、问题二着重分析了不同类型的病人与过敏原检测结果之间的差异性,并将其差异与联系通过可视化的方式表现,同时构建了过敏性疾病的预测模型推断患者可能患何种过敏性疾病,以便提前防范;
- 3、问题三结合选取的环境保护指标——CO<sub>2</sub>浓度,与17种过敏原之间进行了详细的分析,找到其中的联系,对过敏性疾病的前期预防起到一定的作用;
- 4、问题四倡议书部分,针对生态环境破坏与过敏原之间的相互作用关系,对生态环境的保护提出建设性意见以及对过敏性疾病提出预防建议,贴合实际。

### 7.2 模型的缺点

- 1、所给数据只考虑了通过17种过敏原检测结果来确定一个人是否患过敏性疾病,该诊断不够全面,且数据为某医院样本,可能不具有随机性,使得统计分析结果有一定的误差;
- 2、由于时间和专业等原因,搜集的资料可能不够全面,对问题的理解分析上会有所欠缺。

### 7.3 模型的推广

随着我国经济社会的快速发展,生态环境的污染日趋严重,现代人过敏性疾病患病率迅速增加。虽然生态环境污染对人体患病的影响受到社会各界的广

泛关注，但是对各过敏原与环境 and 人体质的科学研究还存在不足。本文通过分析过敏性疾病与季节、年龄之间的关系、过敏原与患者的关系、环境问题与过敏原的关系等，构建 SVM 模型推断患者可能患有何种疾病，为患者选择下一步诊疗方案提供参考，该预测模型可广泛应用于就医前病症预测等方面，为有效预防与降低过敏性疾病提供依据。

## 参考文献

- [1] Boezen H M, van der Zee S C, Postma D S, et al. Effects of ambient air pollution on upper and lower respiratory symptoms and peak expiratory flow in children[J]. The Lancet, 1999, 353(9156): 874-878.
- [2] 马关培, 邹宝兰, 许振成, 等. 广州市某区医院呼吸系统疾病门诊人数与大气污染关系的时间序列研究[J]. 环境与健康杂志, 2012, 29(6): 526-528.
- [3] 朱永生, 梁学友, 钟山, 等. 甲型 H1N1 流感病毒感染与性别关系的统计分析[J]. 中国医药科学, 2015, 5(05): 185-187.
- [4] 李宝红, 孙振球. 对应分析在探讨疾病季节分布中的应用[J]. 中国卫生统计, 2007(06): 593-596.
- [5] 郭磊, 张斌斌, 马瑞娟, 等. 温度对桃离体花药散粉及花粉萌发的影响[J]. 植物生理学报, 2014, 50(3): 269-274.
- [6] 武喆, 刘霞, 张光星. 不同温度对胡萝卜花粉活力的影响[J]. 华北农学报, 2010, 25(04): 116-118.
- [7] Janet Lathrop. Ah-Choo! Expect Higher Grass Pollen and Allergen Exposure in the Coming Century[N/OL]. UMass Amherst. [2014-11-5].  
<http://www.umass.edu/newsoffice/article/ah-choo-expect-higher-grass-pollen-and>

附录

1、过敏原相关化验表及检验标准

SANY

医大学西京医院皮肤科过敏原检测报告单

男

结果

树Ⅱ (桦、栎属) (++) 松属 (+) 大豚草 (++)  
艾蒿 (0) 悬铃木属 (0) 藜属 (0) 禾本科 (++)  
玉属 (++) 树Ⅰ (桉树、榛树、榆科、杨、柳属) (++)  
葎草 (+) 灰藜 (0)  
交链孢菌 (++) 霉Ⅱ (0) 多主枝孢属 (0) 特异靛霉 (0)  
霉Ⅰ (+) 螨Ⅰ (++) 螨Ⅱ (++) 动物毛Ⅰ (++)  
山羊皮毛 (0)  
量检测 鱼 (++) 猪肉 (++) 牛奶 (0) 花生 (++) 鸡蛋 (++)  
桃子 (++) 羊肉 (+) 肉Ⅱ (鸡、鸭、鹅) (++)  
含ASST) 小虾 (0) 蚌类 (+) 红辣椒 (+)

田状胞微粒体抗体 (-)

病 历 号: P34576323-0 床 号:  
性 别:男 年 龄: 3岁3月 费别: 自费 采样时间:  
诊断: 发热原因待查 备 注:

序 代号	项目名称	浓度分级	IU/ml	评价	序 代号	项目名称	浓度分级	IU/ml	评价
1 ts20	柳树/杨树/榆树	1	0.36	弱阳性	16 fs34	鲑鱼	1	0.70	弱阳性
2 w1	普通豚草	0	< 0.35	阴性	17 f24	虾	0	< 0.35	阴性
3 w6	艾蒿	0	< 0.35	阴性	18 f23	蟹	1	0.36	弱阳性
4 ds1	屋尘螨/粉尘螨	0	< 0.35	阴性	19 CCD	CCD	0	< 0.35	阴性
5 h1	屋尘	0	< 0.35	阴性	20 Ind	Control	6	> 100	有效实验
6 e1	猫毛	0	< 0.35	阴性					
7 e2	狗上皮	0	< 0.35	阴性					
8 i6	蟑螂	0	< 0.35	阴性					
9 ms1	点青霉/分枝孢霉	0	< 0.35	阴性					
10 u80	律草	0	< 0.35	阴性					
11 f1	鸡蛋白	0	< 0.35	阴性					
12 f2	牛奶	1	0.36	弱阳性					
13 f13	花生	1	0.36	弱阳性					
14 f14	黄豆	0	< 0.35	阴性					
15 fs33	鲑鱼/龙虾/扇贝	3	5.6	阳性					

申请医生 陆谨 报告时间: 2014-6-13 15:01:24 检验者: 穆局 审核者: 张雯雯  
此结果仅对该标本负责, 且仅供本院医生参考。

年龄: 性别: 科别: 内科 床号: 201010101 上海市第一人民医院		检验依据: 卫生部《全国临床检验操作规程》第三版			建议与解释:	
项目名称	项目代码	结果	提示	单位	参考值	
尘螨/粉尘螨组合		阳性	+		阴性	
豚草/蒿组合		阳性	+		阴性	
动物毛皮屑组合(猫/狗毛皮屑)		阴性			阴性	
蟑螂		阴性			阴性	
霉菌组合(点青霉/分枝孢霉/烟曲霉/交链孢霉)		阴性			阴性	
花粉组合(柏/榆/梧桐/柳/三角叶杨)		阳性	+		阴性	
苜蓿		阴性			阴性	
鸡蛋白/蛋黄组合		阳性	+		阴性	
牛奶		阴性			阴性	
鱼/虾/蟹组合		阳性	+		阴性	
牛肉/羊肉组合		阴性			阴性	
腰果/花生/黄豆组合		阳性	+		阴性	
芒果		阳性	+		阴性	
小麦		阳性	+		阴性	
鱼肝油		阳性	+		阴性	

浓度 (IU/mL)	级别	结果解释
<0.35	0	没有检测到特定抗体
>0.35-0.7	+	轻度过敏, 通常无临床症状但具备一定敏感性。
>0.7-3.5	++	中度过敏, 大量接触后通常会出现临床症状。
>3.5-17.5	+++	高度过敏, 通常临床症状也会出现。
>17.5-50	++++	非常高度的过敏, 通常具有临床症状。
>50-100	+++++	很高程度的过敏

## 2、不同 1(±)数字化值对 5.3 节中的 SVM 模型的拟合结果

1(±)数字化值为 0.5				
'分区'	训练集		测试集	
正确	518	66.01%	229	66.24%
错误	277	33.99%	117	33.76%
总计	785	100%	346	100%
1(±)数字化值为 0.6				
'分区'	训练集		测试集	
正确	354	93.89%	323	93.35%
错误	48	6.11%	23	6.65%
总计	785	100%	346	100%
1(±)数字化值为 0.7				
'分区'	训练集		测试集	

正确	391	49.93%	176	51.13%
错误	394	50.07%	170	48.87%
总计	785	100%	346	100%
1(±)数字化值为 0.8				
'分区'	训练集		测试集	
正确	737	93.89%	323	93.35%
错误	48	6.11%	23	6.65%
总计	785	100%	346	100%
1(±)数字化值为 0.9				
'分区'	训练集		测试集	
正确	551	70.27%	237	68.45%
错误	234	29.33%	109	32.55%
总计	785	100%	346	100%

### 3、基于 SVM 模型的儿科患者可能患疾病类型的预测结果

季节	性别	年 龄 段	树 组 合	普 通 豚草	艾蒿	尘 螨 组合	屋尘	猫毛	狗 上 皮	蟑螂	霉 菌 组合	律草	鸡 蛋 白	牛奶	花生	黄豆	海 鱼 组合1	虾	蟹	疾 病 类型
1	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	1	0	0	2	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1
1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
1	1	1	0	0	3	1	0	0	0	0	0	0	0	1	1	0	0	0	0	1
1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
2	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	1
2	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1
2	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	1	2	0	0	0	2	0	0	2	0	0	1	0	1	1	0	1	0	2
2	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1
2	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1
2	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	1	0	0	0	4	0.8	0	0	0	0	0	0	0	0	0	0.8	0	0	2
4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	1
1	0	1	0	0	0	0	0	0.8	0	0	0	0	0	0	0.8	0	0	0	0	2

1	0	1	0	0	0	0	0	4	0	0	0	0	0	0	4	0	0	0	0	2
1	1	1	0	0	0.8	0	0	2	0	0.8	0	0	0	0	0	0	0	0	0	2
1	1	1	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	2
1	1	1	0	0	0	0	0	2	0.8	0	0	0	0	0	0	0	0	0.8	0	1
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	1
2	1	1	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	1	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	1
2	1	1	0.8	0	0	0.8	0	0	0.8	0	0	0	0	0	0	0	0	0	0	2
3	1	1	0	0	0	0.8	0	4	0	0	0	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0	4	0.8	0	0	0.8	0	0	0.8	0	0	0	0	0	0	2
3	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	1	0	0	0	0.8	0	4	0	0	0	0	0	0	0	0	0	0	0	1
3	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	1	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	1	0	0	0	4	0.8	3	0	0	0	0	0	0	0	0	0.8	0	2	2
3	1	1	0	0	0	4	0.8	0	0	0	0	0	0	0	0	0	0	0	0	2
3	1	1	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	1
3	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2	0	0.8	2
4	0	1	0	0	0	3	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1
4	1	1	0	0	0	4	0.8	0	0	0	0	0	0	0	0	0	0	0.8	0	2
1	1	1	0.8	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2
1	0	1	0	0	0	5	0.8	4	0	0	0	0	0	0	0	0	0	0	0	1



1	1	1	0	0	0	4	2	5	5	0	0	0	0.8	0	0	0	0.8	0	0	2
2	1	1	0	0	0	0	0	0	0	0.8	0	0	0	0	0.8	0	0	0	0	2
2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	2
2	0	1	0	0	0.8	0	0	0	0	0	0	0	0	0.8	0.8	0	0	0	0	1
2	1	1	0	0	0	4	0	0	0	0	0	0	3	2	2	0.8	2	0	0	2
2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	2
2	1	1	0	0	0	4	2	4	2	0	0	0	0	0	2	0	0	0	0	2
2	1	1	0	0.8	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	2
2	1	1	0	0.8	0	0.8	0	0	0	0.8	0	0	0.8	0	2	0	0	0	0.8	2
2	0	1	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	1	0	0	0.8	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	1
3	0	1	0	5	0	0	0	0	0	0	0	3	0	5	0	0	0	0	0	2
3	1	1	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	1
3	0	1	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1
3	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0	0	0.8	0	0	0	0	0	0	2	0	0	0	0	0	1
3	1	1	0	0	0	0	0.8	0	0	0	0.8	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0.8	0.8	2	2	0.8	2	0.8	2	2	5	0	0	2	0	0	1
3	1	1	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	2
3	1	1	0	0	0	0.8	0.8	0	0	0	0	0	0	3	0	0	0	0	0	1
3	1	1	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1

3	1	1	0	0	0	0	0.8	0	0	0.8	0	0	0.8	0	0	0	0	0	0	1
3	0	1	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	1	0	0	0	0	0.8	0	0	0	0	0.8	0	4	0	0	0	0	0	1
3	1	1	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0	0	0.8	0	0	0	0	2	0	5	0	0	0	0	0	1
3	1	1	0	3	0.8	0	0.8	0	0	0	0	4	0	0	0.8	0	0.8	0	0	2
4	0	1	0	0	0	0.8	0	0.8	0	2	0	3	0	0	0.8	0	0	0	0	1
4	0	1	0	0	0	0	0	0.8	0	0	0	4	0	0	0.8	0	0	0	0	2
4	0	1	0	0	0	0	0	0	0	0.8	0	4	0	0	0	0	0	0	0	1
4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	1
4	1	1	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	2
4	1	1	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	2
4	1	1	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	1
4	0	1	0	0	0.8	0	0	0	0	0	0	0.8	0	0	0.8	0	0	0	0	1
4	1	1	0	0	2	0	0	0.8	0	2	0	0	0	0	0	0	0	0	0.8	2
4	1	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2
4	1	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	2
4	1	1	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	1	0	0	0	2	0	0	0	0	0.8	0	0	0	0	0	0	0	0	2
1	1	1	0	0	0	0	0	5	0	0	0	2	0	0	0	0	0	0	0	1
2	1	1	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	2
2	1	1	0	0	0	0.8	0	0	3	0	0	0	0	0	0	0	0	0	0	2

2	1	1	0.8	0	2	5	0	0	0	0	0	0	0.8	0	0	0	0	0	0	2
2	1	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2
3	1	1	0	0	0	5	0	0.8	0	0	0	0	0	0	0	0	0	0	0	2
3	1	1	0	0	0	0	0	0	0	0	0	0	0.8	0	0	2	0	0	0	1
3	1	1	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	1
3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	2
3	1	1	0	0	0	6	2	4	0	0	0	0	0	0	0	0	0	0	0	2
4	0	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	0	1	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	1	1	0	0	0	5	0.8	0.8	0	0	0	0	0	0	0	0	0.8	0	0	1
4	1	1	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
1	1	1	0	5	0	0	0	0	0	0	2	0	0	5	0	0	0	0	0	2
1	1	1	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
1	1	1	0	0	0	0.8	0	0	0	0	0	0	0	4	0	0	0	0	0	1
1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	1
2	0	1	0	0	0	5	0.8	0	0	0.8	0.8	0.8	0	0.8	0.8	0	0.8	0	0	1
2	0	1	0	0	0	0	0	0	0	0	0	0	5	0	0	0	3	0	0	1
2	1	1	0	0	0.8	0	0	0	0	2	0	0	5	0	0	0	0	0	0	1
2	1	1	0	0	0.8	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1
2	1	1	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	1

2	1	1	0	0	0	0	0	0	0	0	0	3	4	0	0	0	5	0	0	2
3	0	1	0	0	0	0	0	0	0	0	0	0	5	0	0	0	4	0	0	1
3	1	1	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
3	1	1	0	0	0	0	0	2	0	0	0	0	6	0	0	0	0.8	0	0.8	1
4	1	1	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	2

#### 4、基于 SVM 模型的体检科与体检中心患者可能患疾病类型的预测结果

季节	性别	年 龄 段	树 组 合	普 通 豚草	艾蒿	尘 螨 组合	屋尘	猫毛	狗 上 皮	蟑螂	霉 菌 组合	律草	鸡 蛋 白	牛奶	花生	黄豆	海 鱼 组合1	虾	蟹	疾 病 类型
2	1	3	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	2
2	1	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	3	3	0	0	0	0	0	0	4	2	0	0	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

1	1	3	2	0	0.8	0	0	0	0	3	0	0	0	2	0	0	0.8	0	0	2
1	1	3	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	3
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	2
1	1	3	0	0	0	0	0	0.8	0	0	0	0	0	2	0	0	0.8	0	0	1
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	0	2	0	0	0	0	0	0	0	0	0	0	0.8	2	0	0	0	0	1	
2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	2	
2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
3	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	2	
3	0	3	0	0	0	0	0.8	0	4	0.8	0	0	0	0	0	0	0.8	3	0.8	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
3	1	2	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	3	0	0.8	2
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
3	0	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	
4	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
4	0	4	0	0.8	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	2	
4	1	4	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	2	
4	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
4	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	

4	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0.8	1
4	1	3	0	0	0	0	0	0.8	0	0	0	0	0	4	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	1
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	3	0	0	0	3	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	3	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	2
1	0	3	0	0	0	0.8	0.8	0	0	0	0	0	0	0	0.8	0	0	0	3	1
2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	3	0	0	0	0	0	2	0	0.8	0	0	0	0	0.8	0	2	0	0	1
2	1	3	0.8	0.8	0	0	4	0	0	0	0	0	0	0	0.8	0	0	0	0	2

2	1	3	0	0.8	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	2
2	1	3	0	0.8	0	0	0	0	0	0.8	0	0	0	0	0.8	0	3	2	3	2
2	0	3	0	0.8	0	0	0	0	0	0.8	0.8	0	0	0	0.8	0	0	0	0.8	1
2	1	3	0	0	0	0	0	0	0	3	0	0	0	0	0.8	0	4	2	2	2
2	0	3	0	0.8	0	2	0	0	0	3	0	0	0	0	0.8	0	4	2	4	2
2	0	2	0	0.8	0	0	0	0	0.8	0	0	0	0.8	0	0.8	0.8	0	0	0	2
3	1	2	0	0	0.8	0	0	0	0	2	0	0	0	2	0	0	0	0	0	2
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	4	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	4	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	1
4	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	1	3	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	1
4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	2	0.8	0	1
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	1
1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	1



1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	4	0	0	0	0	0.8	0	0.8	0	0	0	0	0	0	0.8	0.8	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	2	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0.8	0	0	0	0	0	2	0	0	0	0	0	0	2
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	2	0	0	0	0.8	0	0	3	0	0	0	0	0	0	0	2	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	3	0	0	0	0.8	0	0	2	0	0	0	0	0	0	0	0	0	0	2
2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0	0	0.8	0.8	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	3	0	0	5	0.8	0	0	0	0	0	0	0	0	0	0.8	0	0	2

2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	3	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	1	
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	5	0	0.8	0	0	0	0	0	0	0	0	0	0	3	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
2	1	3	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	1	

2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1
2	0	3	0	0	0	0.8	0	0	0	0	0	0	0	0	0	2	0	0	1
2	0	2	0.8	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	3	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	3
2	1	3	0	0	0	2	0	0	0	0	0	0	0	0	0	3	0	0	3
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0.8	0	0	0	0	0	0	0	0	0	2	0	0	2	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	3	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	2
2	1	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	2	0	0.8	0	0	0	0	0	0	0	0	0	0	0	3
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	2
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2	1	3	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	3
2	0	3	0	0	0.8	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	1

3	0	3	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0.8	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0.8	2
3	1	3	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	2
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	3	0	0	0	0.8	0.8	0	0	0	0	0	0	0	0	0	0	0	0.8	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	4	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	2
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

3	1	3	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0	2
4	1	4	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	0	0	2
4	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
4	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	0	3	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	2	0	0	3	0	0	0	3	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0	0	0	2
1	0	2	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	1
1	0	2	0.8	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0.8	0	0	5	0	0	0	0	2
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	1
3	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	1	2	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	1

## 5、17种过敏原分析源程序

```
C1 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','A2:A83');
C2 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','B2:B139');
C3 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','C2:C96');
C4 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','D2:D355');
C5 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','E2:E208');
C6 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','F2:F134');
C7 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','G2:G126');
C8 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','H2:H148');
C9 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','I2:I113');
C10 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','J2:J115');
C11 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','K2:K124');
C12 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','L2:L135');
C13 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','M2:M220');
C14 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','N2:N66');
C15 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','O2:O151');
C16 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','P2:P112');
C17 = xlsread('/Users/wxy2333/Desktop/learning/数模/2018省赛/A题/数据/1','1','Q2:Q140');

A1 = C1(:);
uniquedata1=unique(A1);
out1=zeros(length(uniquedata1),2);
for ni=1:length(uniquedata1)
    out1(ni,1)=uniquedata1(ni);
    out1(ni,2)=sum(A1==uniquedata1(ni));
end
value_A1 = out1(:,1);
count_A1 = out1(:,2);
x1 = value_A1;
y1 = count_A1;

A2 = C2(:);
uniquedata2=unique(A2);
out2=zeros(length(uniquedata2),2);
for ni=1:length(uniquedata2)
    out2(ni,1)=uniquedata2(ni);
```



```

out2(ni,2)=sum(A2==uniquedatA2(ni));
end
value_A2 = out2(:,1);
count_A2 = out2(:,2);
x2 = value_A2;
y2 = count_A2;

A3 = C3(:);
uniquedatA3=unique(A3);
out3=zeros(length(uniquedatA3),2);
for ni=1:length(uniquedatA3)
out3(ni,1)=uniquedatA3(ni);
out3(ni,2)=sum(A3==uniquedatA3(ni));
end
value_A3 = out3(:,1);
count_A3 = out3(:,2);
x3 = value_A3;
y3 = count_A3;

A4 = C4(:);
uniquedatA4=unique(A4);
out4=zeros(length(uniquedatA4),2);
for ni=1:length(uniquedatA4)
out4(ni,1)=uniquedatA4(ni);
out4(ni,2)=sum(A4==uniquedatA4(ni));
end
value_A4 = out4(:,1);
count_A4 = out4(:,2);
x4 = value_A4;
y4 = count_A4;

A5 = C5(:);
uniquedatA5=unique(A5);
out5=zeros(length(uniquedatA5),2);
for ni=1:length(uniquedatA5)
out5(ni,1)=uniquedatA5(ni);
out5(ni,2)=sum(A5==uniquedatA5(ni));
end
value_A5 = out5(:,1);
count_A5 = out5(:,2);
x5 = value_A5;
y5 = count_A5;

A6 = C6(:);
uniquedatA6=unique(A6);
out6=zeros(length(uniquedatA6),2);
for ni=1:length(uniquedatA6)
out6(ni,1)=uniquedatA6(ni);
out6(ni,2)=sum(A6==uniquedatA6(ni));
end
value_A6 = out6(:,1);
count_A6 = out6(:,2);
x6 = value_A6;
y6 = count_A6;

A7 = C7(:);
uniquedatA7=unique(A7);

```

```

out7=zeros(length(uniqueA7),2);
for ni=1:length(uniqueA7)
out7(ni,1)=uniqueA7(ni);
out7(ni,2)=sum(A7==uniqueA7(ni));
end
value_A7 = out7(:,1);
count_A7 = out7(:,2);
x7 = value_A7;
y7 = count_A7;

A8 = C8(:);
uniqueA8=unique(A8);
out8=zeros(length(uniqueA8),2);
for ni=1:length(uniqueA8)
out8(ni,1)=uniqueA8(ni);
out8(ni,2)=sum(A8==uniqueA8(ni));
end
value_A8 = out8(:,1);
count_A8 = out8(:,2);
x8 = value_A8;
y8 = count_A8;

A9 = C9(:);
uniqueA9=unique(A9);
out9=zeros(length(uniqueA9),2);
for ni=1:length(uniqueA9)
out9(ni,1)=uniqueA9(ni);
out9(ni,2)=sum(A9==uniqueA9(ni));
end
value_A9 = out9(:,1);
count_A9 = out9(:,2);
x9 = value_A9;
y9 = count_A9;

A10 = C10(:);
uniqueA10=unique(A10);
out10=zeros(length(uniqueA10),2);
for ni=1:length(uniqueA10)
out10(ni,1)=uniqueA10(ni);
out10(ni,2)=sum(A10==uniqueA10(ni));
end
value_A10 = out10(:,1);
count_A10 = out10(:,2);
x10 = value_A10;
y10 = count_A10;

A11 = C11(:);
uniqueA11=unique(A11);
out11=zeros(length(uniqueA11),2);
for ni=1:length(uniqueA11)
out11(ni,1)=uniqueA11(ni);
out11(ni,2)=sum(A11==uniqueA11(ni));
end
value_A11 = out11(:,1);
count_A11 = out11(:,2);
x11 = value_A11;
y11 = count_A11;

```

```

A12 = C12(:);
uniquedatA12=unique(A12);
out12=zeros(length(uniquedatA12),2);
for ni=1:length(uniquedatA12)
out12(ni,1)=uniquedatA12(ni);
out12(ni,2)=sum(A12==uniquedatA12(ni));
end
value_A12 = out12(:,1);
count_A12 = out12(:,2);
x12 = value_A12;
y12 = count_A12;

```

```

A13 = C13(:);
uniquedatA13=unique(A13);
out13=zeros(length(uniquedatA13),2);
for ni=1:length(uniquedatA13)
out13(ni,1)=uniquedatA13(ni);
out13(ni,2)=sum(A13==uniquedatA13(ni));
end
value_A13 = out13(:,1);
count_A13 = out13(:,2);
x13 = value_A13;
y13 = count_A13;

```

```

A14 = C14(:);
uniquedatA14=unique(A14);
out14=zeros(length(uniquedatA14),2);
for ni=1:length(uniquedatA14)
out14(ni,1)=uniquedatA14(ni);
out14(ni,2)=sum(A14==uniquedatA14(ni));
end
value_A14 = out14(:,1);
count_A14 = out14(:,2);
x14 = value_A14;
y14 = count_A14;

```

```

A15 = C15(:);
uniquedatA15=unique(A15);
out15=zeros(length(uniquedatA15),2);
for ni=1:length(uniquedatA15)
out15(ni,1)=uniquedatA15(ni);
out15(ni,2)=sum(A15==uniquedatA15(ni));
end
value_A15 = out15(:,1);
count_A15 = out15(:,2);
x15 = value_A15;
y15 = count_A15;

```

```

A16 = C16(:);
uniquedatA16=unique(A16);
out16=zeros(length(uniquedatA16),2);
for ni=1:length(uniquedatA16)
out16(ni,1)=uniquedatA16(ni);
out16(ni,2)=sum(A16==uniquedatA16(ni));
end
value_A16 = out16(:,1);

```

```

count_A16 = out16(:,2);
x16 = value_A16;
y16 = count_A16;

A17 = C17(:);
uniquedatA17=unique(A17);
out17=zeros(length(uniquedatA17),2);
for ni=1:length(uniquedatA17)
out17(ni,1)=uniquedatA17(ni);
out17(ni,2)=sum(A17==uniquedatA17(ni));
end
value_A17 = out17(:,1);
count_A17 = out17(:,2);
x17 = value_A17;
y17 = count_A17;

figure()
width = 400;
height = 300;
subplot(6,3,1);
plot(x1,y1);
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});
xlabel('(a)树组合');
ylabel('过敏人数');

subplot(6,3,2);
plot(x2,y2);
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});
xlabel('(b)普通豚草');
ylabel('过敏人数');

subplot(6,3,3);
plot(x3,y3);
xlabel('(c)艾蒿');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,4);
plot(x4,y4);
xlabel('(d)尘螨组合');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,5);
plot(x5,y5);
xlabel('(e)屋尘');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

```

```

subplot(6,3,6);
plot(x6,y6);
xlabel('(f)猫毛');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,7);
plot(x7,y7);
xlabel('(g)狗上皮');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,8);
plot(x8,y8);
xlabel('(h)蟑螂');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,9);
plot(x9,y9);
xlabel('(i)霉菌组合');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,10);
plot(x10,y10);
xlabel('(j)律草');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,11);
plot(x11,y11);
xlabel('(k)鸡蛋白');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,12);
plot(x12,y12);
xlabel('(l)牛奶');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,13);
plot(x13,y13);

```

```

xlabel('(m)花生');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,14);
plot(x14,y14);
xlabel('(n)黄豆');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,15);
plot(x15,y15);
xlabel('(o)海鱼组合');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,16);
plot(x16,y16);
xlabel('(p)虾');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

subplot(6,3,17);
plot(x17,y17);
xlabel('(q)蟹');
ylabel('过敏人数');
set(gca,'XTick',1:1:4);
set(gca,'XTickLabel',{'春','夏','秋','冬'});

```

## 6、.过敏性疾病随时间变化趋势源程序

```

A = [];

A1 = A(:,1);
A2 = A(:,2);
A3 = A(:,3);
A4 = A(:,4);
B1 = tabulate(A1);
B2 = tabulate(A2);
B3 = tabulate(A3);
B4 = tabulate(A4);

value_A1 = B1(:,1);
count_A1 = B1(:,2);
percent_A1 = B1(:,3);
x1 = value_A1;
y1 = count_A1;
z1 = percent_A1;
value_A2 = B2(:,1);

```

```

count_A2 = B2(:,2);
percent_A2 = B2(:,3);
x2 = value_A2;
y2 = count_A2;
z2 = percent_A2;
value_A3 = B3(:,1);
count_A3 = B3(:,2);
percent_A3 = B3(:,3);
x3 = value_A3;
y3 = count_A3;
z3 = percent_A3;
value_A4 = B4(:,1);
count_A4 = B4(:,2);
percent_A4 = B4(:,3);
x4 = value_A4;
y4 = count_A4;
z4 = percent_A4;

A1 = A1(:);
uniquedata1=unique(A1);
out1=zeros(length(uniquedata1),2);
for ni=1:length(uniquedata1)
out1(ni,1)=uniquedata1(ni);
out1(ni,2)=sum(A1==uniquedata1(ni));
end
value_A1 = out1(:,1);
count_A1 = out1(:,2);
x1 = value_A1;
y1 = count_A1;

A2 = A2(:);
uniquedata2=unique(A2);
out2=zeros(length(uniquedata2),2);
for ni=1:length(uniquedata2)
out2(ni,1)=uniquedata2(ni);
out2(ni,2)=sum(A2==uniquedata2(ni));
end
value_A2 = out2(:,1);
count_A2 = out2(:,2);
x2 = value_A2;
y2 = count_A2;

A3 = A3(:);
uniquedata3=unique(A3);
out3=zeros(length(uniquedata3),2);
for ni=1:length(uniquedata3)
out3(ni,1)=uniquedata3(ni);
out3(ni,2)=sum(A3==uniquedata3(ni));
end
value_A3 = out3(:,1);
count_A3 = out3(:,2);
x3 = value_A3;
y3 = count_A3;

A4 = A4(:);
uniquedata4=unique(A4);
out4=zeros(length(uniquedata4),2);

```

```

for ni=1:length(uniquedata4)
out4(ni,1)=uniquedata4(ni);
out4(ni,2)=sum(A4==uniquedata4(ni));
end
value_A4 = out4(:,1);
count_A4 = out4(:,2);
x4 = value_A4;
y4 = count_A4;

plot(x1,y1,'-r');
hold on
plot(x2,y2,'-b');
hold on
plot(x3,y3,'-y');
hold on
plot(x4,y4,'-g');
set(gca,'XTick',1:2:99);
set(gca,'XTickLabel',{'12年冬','13年春','13年夏','13年秋','13年冬','14年春',
'14年夏','14年秋','14年冬','15年春','15年夏','15年秋','15年冬','16年春','16
年夏','16年秋','16年冬','17年春','17年夏','17年秋','17年冬'})

```