

2017 深圳杯研究论文

深圳市蔬菜安全风险评估与抽样设计



Nankai University, Tianjin

姓 名 _____ 庞敏行 _____
姓 名 _____ 王真 _____
姓 名 _____ 王雅晴 _____
选 择 题 目 _____ D _____
日 期 _____ 2017 年 8 月 _____

摘 要

新鲜蔬菜在中国居民的膳食中占据很大的比例,但是近年来蔬菜安全领域却危机频现。为了有效地评估从而防控新鲜蔬菜的多发风险,制定出一套针对市场在售蔬菜的科学可行的抽检方案,便具有十分重要的意义。本文以深圳市新鲜蔬菜为研究对象,在分析食品安全监管等部门现有数据的基础上,结合收集到的多方数据,评估忽略某些输入渠道造成的蔬菜安全风险,建立一个评估新鲜蔬菜食品安全风险的模型,并设计出一套改进现有方案的抽检方案。

对于被忽略的某些输入渠道造成的风险,本文建立了渠道安全风险分析模型,运用灵敏度分析得出忽略这些渠道所导致的食物安全风险。在建立渠道模型时,本文首先基于各输入渠道的属性指标确立一套风险评估体系,接着引入熵权法求得各渠道属性指标的熵值,从而确定各属性指标的权重。最后通过基于灰色关联系数的 TOPSIS 法求出各渠道与正负理想渠道的贴近度,从而确定各渠道的安全系数。

对于深圳市新鲜蔬菜安全风险的评估,本文建立了新鲜蔬菜安全风险评估模型。基于影响新鲜蔬菜安全的多方面因素建立一套评估新鲜蔬菜安全风险的指标体系,并通过层次分析法和熵权法分别确定一级和二级指标的权重。在这里,本文将蔬菜按官方八大类进行分类,分别求出各类蔬菜的风险系数后,通过加权求和得到全市新鲜蔬菜整体的安全风险系数。

对于更优抽检方案的设计,考虑到传统抽检方案忽略了新兴的输入渠道、没有按抽检结果动态调整抽检方案等不足之处,本文建立了一套新的抽检方案。通过对已有抽检结果的数据统计,按照当月实际或预测的销售情况得出较贴近实际抽检结果的抽检方案,即按照每个季节不同渠道和种类分配抽样比例,再根据风险系数对每类蔬菜和渠道的抽样量进行调整。最后采用蒙特卡罗方法对抽检方案进行检验,即利用线性同余发生器设置均值和方差,其产生的伪随机序列即可模拟蔬菜的食品安全情况,从而模拟实际抽检操作,检验抽检方案的可行性及准确度。

关键词: 蔬菜安全; 安全风险系数; TOPSIS 法; 蒙特卡罗方法

一 问题重述

1.1 问题背景

新鲜蔬菜的安全是居民食品安全领域中十分重要的一环。随着社会经济的发展和居民生活水平的提高,城镇所销售的新鲜蔬菜的输入渠道越来越具有多样性,居民消费的蔬菜种类也越来越丰富。与之伴随而来的则是越来越频发的新鲜蔬菜安全风险问题,以及相关监管部门在抽检监督等工作中的困难。在我国,食品的安全风险评估起步较晚,因而无论是有关法律制度的制定,还是风险管理机制的运行都存在不少瑕疵,需要作进一步探讨。

1.2 提出问题

对蔬菜安全风险的监测除了要进行精准的风险评估,更为重要的是实施有效的抽检工作。随着经济发展多元化趋势日渐明显,我国的蔬菜消费市场正处在急速转型的时期,新兴输入渠道的兴起,新种类蔬菜的培育和消费等新形势都对传统的蔬菜安全管理和抽检方法提出了新的要求。尽管不同的风险评估方法和抽检方案都在不断地进行理论上的创新,以解决实际操作中遇到的各种问题,但遗憾的是,这一领域仍然存在一些无法完全解决的难题,比如对进口蔬菜的安全检测标准的统一等问题。这些问题都需要我们在建立模型时,结合其他学科的知识进行多方面的考虑。

本文旨在基于可收集到的关于深圳市新鲜蔬菜输入与消费的多方数据,建立数学模型,评估忽略某些新兴输入渠道给新鲜蔬菜食品安全造成的风险;进一步利用现有数据结合模型评估深圳市新鲜蔬菜的安全风险;最后,基于以上研究成果,分析有关部门现行抽检制度存在的弊端,设计出一套改进的抽检方案;从而能够给有关部门在新鲜蔬菜安全的监管与防控方面提出合理建议。

二 问题分析

针对前文提出的新鲜蔬菜安全风险领域存在的难题,本文主要通过建立渠道风险模型、蔬菜风险评估模型和抽检模型寻求解决方案,主要思路如下:

对于需要收集的数据,经过整理之后在附件 A 部分。

对于评估抽检中忽略某些渠道造成的食品安全风险,考虑建立渠道风险评估模型,用能够反映不同种类输入渠道的属性的几个指标,量化地描述各类渠道,而后用熵权法确定各属性的权重,再引入灰色关联度分析法和 TOPSIS 法评估各输入渠道存在的安全风险。最后用灵敏度分析的方法计算忽略某些渠道所增加的风险。

对于深圳新鲜蔬菜安全风险的评估,考虑蔬菜风险评估模型,构建多级指标评估体系,用影响蔬菜安全风险的因素对应的指标,按照蔬菜类别对其存在的安全风险进行评估,并采用层次分析法确定一级指标的权重。将不同种类蔬菜的风险评估数据进行汇总后得到深圳市新鲜蔬菜整体的安全风险。

对于抽检方案的设计,考虑使得抽检方案既能反映现有蔬菜存在的安全问题,又能起到预警潜在蔬菜安全风险的作用,通过构造与时间属性相关和无关的两个函数,确定抽样比例,从而确定抽检方案。并采用蒙特卡罗方法对抽检计划进行检验。

最后,本文拟在科学模型的基础上,从问题研究者的角度对深圳市新鲜蔬菜食品安全监管工作提出一些浅显的建议。

三 假设与符号

3.1 模型假设

1. 蔬菜质量呈正态分布
2. 同一种类的蔬菜具有一致性
3. 已知较为准确的市场质量函数
4. 每个批次抽样的总数一定
5. 在稳定时期,短期内蔬菜合格率波动不大
6. 根据调查,每个渠道中蔬菜种类的比例基本一致

3.2 符号约定

r_{ij}	第 i 种渠道第 j 个属性的归一化数据
t_i	第 i 个渠道的蔬菜销量比
ω_j	各渠道第 j 类属性的权重
C_{0i}^+	第 i 种渠道的安全性
γ	渠道总风险系数
m_i	第 i 个蔬菜安全属性的权重
d_i	第 i 种危险源的毒性指数
δ_{ij}	第 i 种危险源在第 j 类蔬菜中的含量超额幅度 (处理后)
n_{ij}	第 i 种危险源在第 j 类蔬菜中的出现频次
α_{1j}	第 j 类蔬菜在危险源含量超额幅度上的风险系数
α_{2j}	第 j 类蔬菜在危险源出现频次上的风险系数
α_i	第 i 类蔬菜在危险源方面的风险系数
s_{ij}	第 i 类蔬菜在第 j 个季节的销量比
ω_{ij}	第 i 类蔬菜在第 j 个季节出现的区域数
β_i	第 i 类蔬菜在数量安全方面的风险系数
Q_i	第 i 类蔬菜处理后的不合格率
f_i	第 i 种渠道的风险系数
S_{ij}	第 i 类蔬菜的第 j 种渠道的销量比
M_i	第 i 类蔬菜的风险系数
s_i	第 i 类蔬菜的销量比
w	函数 g_i 的权重
u	函数 f_i 中关于均值的权重
v	函数 f_i 中关于标准差的权重

四 模型建立与求解

4.1 问题二的分析与求解

4.1.1 模型分析

- 由问题要求,建立渠道风险评估模型,得出每种输入渠道的风险系数之比。通过灵敏度分析得出忽略某些渠道导致的安全风险对深圳市新鲜蔬菜整体安全风险水平的影响。
- 将蔬菜的输入渠道归为五类,即批发市场,农贸市场,超市及零售商店,网络营销以及点对点直接配送。
- 将每种输入渠道用蔬菜流量,监察力度,行业素质,流通环节数这四个属性进行描述。每一输出渠道的食品安全风险由这四个渠道属性决定。

4.1.2 模型建立

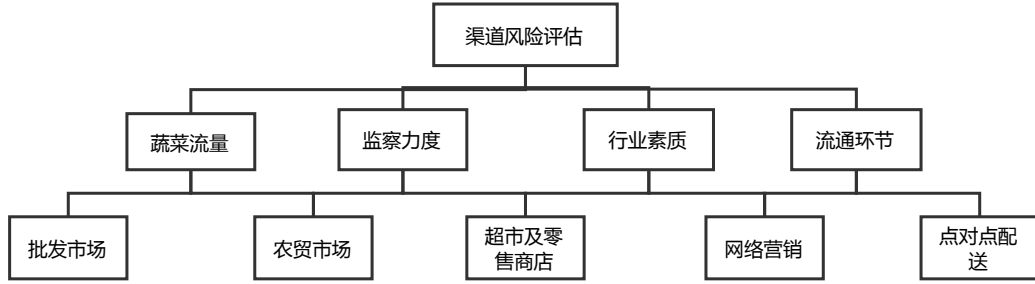


图 1: 模型示意图

定性指标量化

专家对各定性指标因素集给出了语言真值(评价结果),其无法和定量进行对比,因此需对定性指标量化,量化方法是将指标分成9个等级,量化定性指标:指标等级的评分由高到低为:9,8,7,6,5,4,3,2,1.评价等级划分采取极好,很好,好,较好,一般,较差,差,很差,极差。

用熵权法确定各指标的权重(属性)

1. 归范化处理:在上文构造的渠道风险评估模型中存在多个指标,这些指标的量纲是不同的,因此在处理前需要对其先做归范化处理。这里,我们采用极差转换法。对于效益性指标

$$r'_i = \frac{r_i - \min r_i}{\max r_i - \min r_i}, (i = 1, 2, \dots, n) \quad (1)$$

对于成本性指标。

$$e'_j = \frac{\max e_j - e_j}{\max e_j - \min e_j}, (j = 1, 2, \dots, n) \quad (2)$$

2. 确定各指标的权重系数

归范化处理后的标准矩阵为:

$$R = (r_{ij})_{m \times n} = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix} \quad (3)$$

熵值为:

$$x_j = -\frac{1}{\ln m} \sum_{i=1}^m \ln(r_{ij}), 1 \leq i \leq m, 1 \leq j \leq n. \quad (4)$$

各属性的权重系数为:

$$\omega_j = \frac{1 - x_j}{n - \sum_{j=1}^n x_j}, 1 \leq j \leq n. \quad (5)$$

基于熵权的评价矩阵构建

代入各属性的权重系数 ω_j , 将规范化后的数据加权, 构建加权规范化决策矩阵 Y

$$Y = (y_{ij})_{m \times n} = \begin{pmatrix} r_{11} \cdot \omega_1 & \cdots & r_{1n} \cdot \omega_n \\ \vdots & \ddots & \vdots \\ r_{m1} \cdot \omega_1 & \cdots & r_{mn} \cdot \omega_n \end{pmatrix} \quad (6)$$

用基于灰色关联度的 TOPSIS 法评价各渠道风险

1. 构建加权规范化矩阵的正理想渠道和负理想渠道, 记为 C^+ , C^- 。正理想解即为所有属性都由于已知渠道的理想渠道。负理想渠道为所有属性都劣与已知渠道的理想渠道。

$$C^+ = (y_{01}^+, \cdots, y_{04}^+) = (\min_i \{y_{i1}\}, \max_i \{y_{i2}\}, \max_i \{y_{i3}\}, \min_i \{y_{i4}\}). \quad (7)$$

$$C^- = (y_{01}^-, \cdots, y_{04}^-) = (\max_i \{y_{i1}\}, \min_i \{y_{i2}\}, \min_i \{y_{i3}\}, \max_i \{y_{i4}\}). \quad (8)$$

2. 构造灰色关联系数矩阵。计算第 i 个渠道与正理想渠道关于第 j 个指标的灰色关联系数 ϵ_{ij}^+ 。从而建立各渠道与正理想渠道的灰色关联系数矩阵 $R^+ = (\epsilon_{ij}^+)_{m \times n}$ 。同理, 建立与负理想解的灰色关联系数矩阵 $R^- = (\epsilon_{ij}^-)_{m \times n}$ 。

其中, 灰色关联系数为

$$\epsilon_{ij}^+ = \frac{\min \min |y_{0j}^+ - y_{ij}| + p \max \max |y_{0j}^+ - y_{ij}|}{|y_{0j}^+ - y_{ij}| + p \max \max |y_{0j}^+ - y_{ij}|} \quad (p \text{ 通常取 } 0.5) \quad (9)$$

3. 计算各渠道与正负理想渠道的灰色关联度 r_i^+ 和 r_i^- 。

$$r_i^+ = \frac{1}{4} \sum_{j=1}^4 \epsilon_{ij}^+, \quad r_i^- = \frac{1}{4} \sum_{j=1}^4 \epsilon_{ij}^- \quad (10)$$

4. 构造各渠道的灰色关联贴近度 C_{0i}^+ 。

$$C_{0i}^+ = \frac{r_i^-}{r_i^+ + r_i^-}, (i = 1, \cdots, 5) \quad (11)$$

各渠道与负理想渠道相对贴近度之比即为各渠道的安全性之比。因为 C_{0i}^+ 是各渠道与负理想渠道的灰色关联贴近度, 所以 C_{0i}^+ 的值越大, 渠道的风险越高。

4.1.3 模型求解

1. 通过 MATLAB 求出各属性的熵值 (具体算法附后), 结果为

	蔬菜流量	监察力度	行业素质	流通环节
熵值	0.2283	0.1931	0.2948	0.2838

2. 运用 MATLAB 求各渠道分别于正负理想渠道的关联度(具体算法附后),得到关联度矩阵:
与正理想渠道的关联系数矩阵:

$$\begin{pmatrix} 0.8451 & 0.6209 & 0.3611 & 0.6739 \\ 1 & 0.6756 & 0.3986 & 0.9154 \\ 0.8374 & 0.7409 & 0.5033 & 0.8731 \\ 0.5775 & 0.4689 & 0.3986 & 0.8731 \\ 0.5609 & 0.4995 & 0.3611 & 0.6739 \end{pmatrix} \quad (12)$$

与负理想渠道的关联系数矩阵:

$$\begin{pmatrix} 0.3413 & 0.9924 & 0.9278 & 1 \\ 0.3429 & 0.9830 & 0.9457 & 0.9884 \\ 0.3412 & 0.9738 & 0.9815 & 0.9609 \\ 0.3367 & 0.9864 & 0.9457 & 0.9609 \\ 0.3362 & 0.9959 & 0.9278 & 1 \end{pmatrix} \quad (13)$$

3. 各渠道的灰色关联贴近度即为:

	批发市场	农贸市场	超市及零售商店	网络营销	点对点配送
关联度	0.574	0.778	1.09	1.04	0.525

4.1.4 运用灵敏度分析考虑忽略渠道的安全风险影响

渠道总风险系数 γ 即可表示为,

$$\gamma = \sum C_{0i} \cdot t_i \quad (14)$$

通过比较只考虑已知渠道 ($i = 1, 2, 3$) 的风险系数 γ_1 和综合考虑全部渠道 ($i = 1, 2, 3, 4, 5$) 的风险系数 γ_0 , 即可估算出忽略某些渠道带来的风险。其中,

$$\gamma_1 = \sum_{i=1}^3 C_{0i} \cdot t_i \quad (15)$$

$$\gamma_0 = \sum_{i=1}^5 C_{0i} \cdot t_i \quad (16)$$

则忽略某些渠道所增加的风险比率为:

$$D = \frac{\gamma_0 - \gamma_1}{\gamma_0} = 0.0265 \quad (17)$$

进一步考虑, 利用安全风险评估模型进行评估更具有针对性, 能体现出忽略渠道对于蔬菜安全风险评估的影响。由于安全风险模型经过了检验, 具有一定的可靠性。使用灵敏度分析的思想, 主要方法为

- 利用近一年的数据通过渠道分析模型计算出各数菜的危险程度, 并代入安全风险评估模型计算出当年的蔬菜安全风险系数。
- 把网络销售和点对点配送的数据忽略, 重新计算出蔬菜的危险程度和同年的蔬菜安全风险系数。
- 对比分析两次的蔬菜安全风险系数的变化对于蔬菜安全风险等级评估的影响。

4.1.5 结果分析

经计算,原蔬菜安全风险系数¹为 $A_0 = 1.60$, 忽略了两种新兴渠道得到的安全系数为 $A_1 = 1.36$, 致使模型评估的危险等级降低。由于新兴渠道目前的规模还在发展阶段, 对于食品安全风险的影响还不是很大, 但是也不能忽略, 这一点和实际相符。由于新兴渠道的监管不成熟, 危险系数也会较大, 点对点配送在模型中的危险系数为 1.04, 处于较高的水平, 也和实际相符合。

结合针对渠道的和蔬菜整体安全风险的分析, 目前忽略一些新兴渠道带来的风险影响在 10% 以内。但是受季节的影响, 网络营销渠道的销量极不稳定, 风险具有爆发性的特点, 容易在短期内对食品安全造成较大的影响。考虑到这一点, 常规抽检中忽略的渠道风险不容忽视。

4.2 问题三的分析与求解

4.2.1 模型分析

- 将蔬菜按抽检计划分为八大类, 评估蔬菜整体的安全风险。
- 将蔬菜的安全风险用危险源, 市场 (蔬菜数量安全), 不合格率以及渠道风险性四个方面的指标描述。危险源指标又分为危险源的危害等级, 含量超额幅度以及出现的频次三个二级指标进行描述。蔬菜数量安全分为销量和销售范围 (出现在各行政区的计数) 两个三级指标。
- 得出整体蔬菜风险评估模型中各个指标的权重, 并代入已得抽检结果得出深圳市新鲜蔬菜整体风险系数。

4.2.2 模型建立

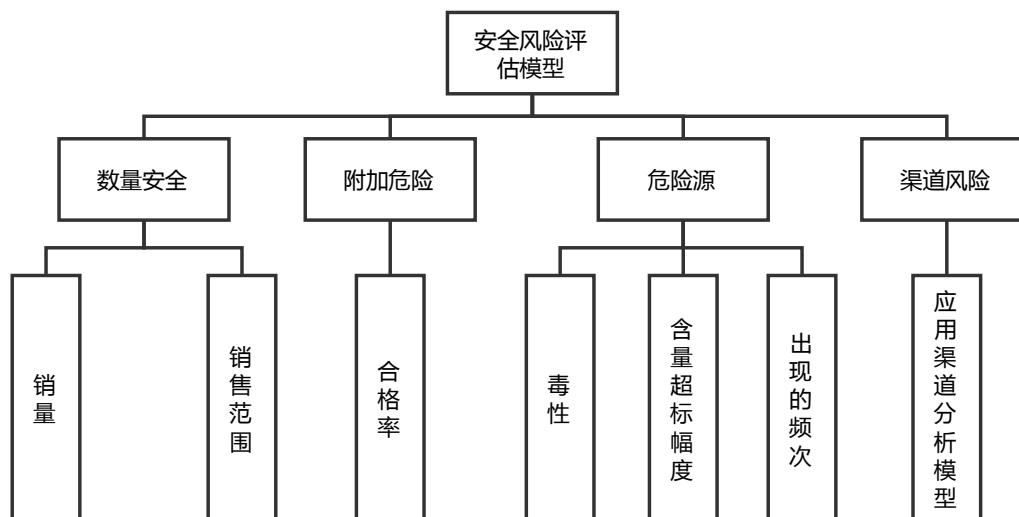


图 2: 模型示意图

用层次分析法确定一级指标的权重

将四个一级指标 A_1, A_2, A_3, A_4 对蔬菜 O 的影响两两作比较。用 a_{ij} 表示 A_i 和 A_j 对 O 的影响之比, 全部比较结果可用成对比较矩阵

$$A = a_{ij}, \quad a_{ij} > 0, \quad a_{ij} = \frac{1}{a_{ji}} \quad (18)$$

¹ 详见安全风险评 估模型

表示。比较尺度用 1-9 尺度。如果比较矩阵通过一致性检验,则可用 A 最大特征根的特征向量归一化后作为权向量 m_i 。

确定危险源下属二级指标对应的权向量

1. 抽象指标具体化

对象为每一类蔬菜,对于危险源的危害等级,用危险源(农药)的毒性²描述,标准为 LD50,大鼠(雄)。(即雄性大鼠的半数致死量),第 i 种危险源的 LD50³为 p_i 。按照中国农药网依照 LD50 值对危险源毒性的分类,有微毒、低毒、中毒、高毒、剧毒五种情况。本文在量化该抽象指标时,分别用 1、3、5、7、9 对应;对于含量超额幅度,用平均检测量对于规定上限的超出百分比表示,第 i 种危险源的含量超额幅度为

$$\Delta_{ij} = \frac{d_{ij} - D_{ij}}{D_{ij}} \quad (19)$$

危险源的出现频次即用计数法表示为 n_{ij} 。

由于各危险源的含量超额幅度差异较大,为确保整体数据统计的合理性,我们超额幅度取对数作为初始数据。 $\delta_{ij} = \ln \Delta_{ij}$

2. 数据处理

在本模型中,我们一直以不同类别的蔬菜作为研究对象,并且同一危险源在不同类蔬菜中出现的概率不同。为准确评估不同类蔬菜的风险,本文将作为危险源固有属性的毒性乘到其它两个属性的数据上,以表明危险源施加在不同类蔬菜上的不同影响。具体操作如下:

(1) 考虑到含量超额幅度和危险源频次的量纲不同,直接加权影响风险系数评估的准确性。本文将 δ_{ij} 和 n_{ij} 分别除以常数 δ_{0j} 和 n_{0j} 。得到 δ_{ij}^1 和 n_{ij}^1 。其中 δ_{0j} 和 n_{0j} 大致与 $\sum_{j=1}^{j=7} \delta_{0j}$ 和 $\sum_{j=1}^{j=7} n_{0j}$ 相同。(相当于按比例放缩至大致和为 1)

(2) 将危险源的毒性指数 d_i 分别与 δ_{ij}^1 和 n_{ij}^1 相乘,并将每类蔬菜中不同危险源的同一属性的数据求和既得该类蔬菜在该属性的风险系数。即就是

$$\alpha_{1j} = \sum_{i=1}^{14} d_i \times \delta_{ij} \quad (20)$$

$$\alpha_{2j} = \sum_{i=1}^{14} d_i \times n_{ij} \quad (21)$$

3. 用 Delphi 法确定权重. 我们与进行鲜蔬抽检工作的相关专家进行详细咨询,确定含量超额幅度和出现频次的权重。

4. 加权得到该类蔬菜的危险源风险系数 α_i

用熵权法确定蔬菜的数量安全下属二级指标对应的权向量

设第 i 类蔬菜第 j 个季节的销量比为 s_{ij} , 第 i 类蔬菜第 j 个季节蔬菜不合格情况出现的区域数为 w_{ij} 则可列出两组数据求出该类蔬菜每年在这两方面的平均水平

$$\bar{s}_i = \frac{\sum_{j=1}^4 s_{ij}}{4} \quad (22)$$

$$\bar{w}_i = \frac{\sum_{j=1}^4 w_{ij}}{4} \quad (23)$$

²毒性等级由中国农药毒性分级标准(共五级)

³LD50 均用对大鼠(雄)急性经口(原药)的值;若是一个取值范围则用其中值

表 1: 各类危险源 LD50 值

危险源	LD50(毫克/千克)	毒性等级
阿维菌素	10	高毒
敌敌畏	80	中毒
啉虫脒	217	中毒
毒死蜱	82	高毒
多菌灵	15000	微毒
氟虫腈	97	高毒
甲胺磷	25	剧毒
甲拌磷	3.7	剧毒
甲基异柳磷	21	剧毒
呋喃丹	11	高毒
联苯菊酯	54.5	中毒
溴氰菊酯	136.5	中毒
氧乐果	500	中毒
乙酰甲胺磷	945	低毒

通过加权求和可得每类蔬菜的在数量安全方面的危险系数 β_i

蔬菜的附加危险系数

使用不合格率 q_i 进行归一化处理得到 Q_i

确定渠道风险系数

利用渠道分析模型对每个渠道的危险系数 f_i , 可计算该类蔬菜的渠道风险系数

$$\eta_i = \sum_{j=1}^5 f_j \times S_{ij} \quad (24)$$

求出该类蔬菜风险系数

1. 考虑到四个风险系数 $\alpha_i, \beta_i, Q_i, \eta_i$ 的量纲不同, 数量级差异过大影响总风险系数的评估。因此我们将每组数列 $(\alpha_1, \alpha_2, \dots, \alpha_7), (\beta_1, \beta_2, \dots, \beta_7), (Q_1, Q_2, \dots, Q_7), (\eta_1, \eta_2, \dots, \eta_7)$ (由于水生类蔬菜在附件中没有相关数据, 因此本文按七类蔬菜来计算) 分别除以常数 $\alpha_0, \beta_0, Q_0, \eta_0$, 使得每个一级指标的风险系数对总风险系数的影响大致相同。其中, 常数 $\alpha_0, \beta_0, Q_0, \eta_0$ 是通过大量深圳市蔬菜抽检结果的数据所确定的, 它们的值分别与 $\sum \alpha_i, \sum \beta_i, \sum Q_i, \sum \eta_i$ 的值大致相同。(相当于按比例放缩至大致和为 1)

2. 代入各权重求和得该类蔬菜的风险 M_i

求出整体蔬菜的安全系数

直接用每类蔬菜的销量 s_i 对该类蔬菜的风险 M_i 加权求和, 即

$$A = \sum_{i=1}^8 s_i \times M_i \quad (25)$$

确定风险系数评判标准

为了使求得结果更加直观, 我们希望建立一套与风险系数对应的语言性数据的评判标准。因此本文运

用边界数据法, 根据实际情况创设两套最安全和最危险的蔬菜销售市场的原始数据, 视为边界数据。将它们代入渠道模型和安全风险模型得出的风险系数即为风险系数的两个边界值, 对应的语言性数据即为最安全和最危险, 分别记为 e_1, e_2 。

将区间 $[e_2, e_1]$ 等分为五个区间, 分别为危险, 较危险, 一般, 较安全, 安全。判断算得的整体蔬菜风险系数 A 的对应区间即可评估目前深圳市整体蔬菜安全情况。

4.2.3 模型求解

1. 通过层次分析法求出一级指标的权重(程序附后)

表 2: 层次分析法

一级指标	数量安全	附加危险	危险源	渠道风险	权重	CR=CI/RI(RI=0.8931)
数量安全	1	5	6	8	0.638566483	
附加危险	0.2	1	3	5	0.213517572	
危险源	0.16666667	0.33333333	1	2	0.093668309	0.055513469
渠道风险	0.125	0.2	0.5	1	0.054247635	符合 CR<0.1

2. 确定每类蔬菜各属性的风险系数

表 3: 风险系数求解

	数量安全	不合格	危险源	渠道安全
豆类	0.10933816	0.1483871	0.15961382	0.06865444
根茎类	0.08589179	0.0308642	0.04933117	0.07019944
瓜类	0.1495971	0.01907357	0.06043435	0.16364333
茄果类	0.09523865	0.01754386	0	0.09847278
鳞茎类	0.00215652	0	0.0366892	0.00312444
叶菜类	0.34843671	0.03902954	0.23499153	0.43139556
芸薹属类	0.19909758	0.08883249	0.20439736	0.17431222

3. 确定每一类蔬菜的风险系数和整体蔬菜的总系数

表 4: 风险系数求解

	豆类	根茎类	瓜类	茄果类	鳞茎类	叶菜类	芸薹属类
风险系数	0.27913895	0.13258871	0.19979169	0.10040342	0.02948467	0.53370403	0.38051904
总系数							1.655630514

4. 确定蔬菜风险系数评定标准

表 5: 安全边界

	数量安全	不合格	危险源	渠道安全	总计
豆类	0.10933816	0.1483871	0.15961382	0.06865444	0.18101281
根茎类	0.036682844	0.015432099	0.062618636	0.077079778	0.088387846
瓜类	0.089654583	0.009536785	0.045002059	0.175408	0.125621202
茄果类	0.048221526	0.00877193	0	0.107091778	0.050797745
鳞茎类	0.001341707	0	0.033259979	0.003776444	0.026121862
叶菜类	0.20783329	0.019514768	0.215717256	0.437447556	0.373342554
芸薹属类	0.111230374	0.044416244	0.164588722	0.181232889	0.248330559
					1.093614578

表 6: 危险边界

	数量安全	不合格	危险源	渠道安全	总计
豆类	0.141153294	0.222580645	0.310261985	0.063691111	0.448897721
根茎类	0.12655438	0.046296296	0.233536493	0.064904444	0.31555324
瓜类	0.20939195	0.028610354	0.072339196	0.149979333	0.271186463
茄果类	0.150705109	0.026315789	0.049970781	0.091474444	0.19562298
鳞茎类	0.036979729	0.01534	0.083765916	0.002952444	0.104413853
叶菜类	0.445560956	0.058544304	0.515061934	0.380196222	0.845444113
芸薹属类	0.244841563	0.133248731	0.29373571	0.156048222	0.507938058
					2.689056429

蔬菜风险评判标准则为

风险系数	1.09-1.41	1.41-1.73	1.73-2.05	2.05-2.37	2.37-2.69
	安全	较安全	一般	较危险	危险

深圳市整体蔬菜风险系数为 1.66,属于较安全范畴。

4.3 问题四的分析与求解

根据食药监局专家的意见和对实际需要对分析,一个好的抽样模型应该是使得结果既能反映出整体的蔬菜质量状况,又能体现出某些风险状况。仅仅为了提高抽检结果和实际市场状况的吻合度或仅仅提高不合格产品的检出率都是与实际需要不符合的。根据这一思路,我们结合蒙特卡罗方法对思想,建立的抽样模型主要思路为:在总的抽检数量恒定的情况下,根据蔬菜的销售安全性(涉及安全系数,前一期合格率及其方差)对销量进行加权,确定抽样比例从而确定抽样方案。

4.3.1 种类分析

根据统计学的抽样原理为了提高抽检的效率,本模型根据实际的数据和抽检的目标综合考虑该类销量相关的函数 g_i ,和前几个月抽检情况相关的函数 f_i ,所以对于每一类蔬菜的抽检量的权重为

$$T_i = \phi(u, v, w)$$

固定部分 g_i

- 运用历年当季各种蔬菜销量比重的数据预测此月份的比重,记为 ξ_i ,权重为 w 。

时间相关部分 f_i

为了更精确地反映出一种不合格蔬菜的风险的水平和趋势,我们综合考虑上个月不合格情况的两个指标

- 上一期抽查出现不合格状况的均值 μ_i ,权重为 u 。
- 上一期该种类出现不合格状况的标准差 β_i ,权重为 v 。

标准化处理后得到关于均值和方差的标准正态分布 $norm(\mu_i, \beta_i)$, 利用计算机分别求解标准正态曲线在 1 右侧的面积,记为 f_i , 面积越大则说明风险越大。

故每一类蔬菜的抽样比例为

$$q_i = \frac{f_i}{\sum_{j=1}^8 f_j} \quad (26)$$

最后,每类蔬菜的抽样比例为

$$K = g_i \times f_i$$

4.3.2 渠道分析

方法与种类分析类似,但是考虑到各个渠道的安全风险不同,且新兴渠道与传统渠道的差异较大,我们对目标函数增加一个权,代表该渠道的安全风险。适用于深圳市和各辖区的抽检计划的确定。目前的常规抽检中缺少点对点配送和网络营销渠道的抽检,此模型可应用于这两方面的分析。

- 运用渠道分析模型中得到的安全系数 γ ,权重记为 p 。
- 运用历年当季各渠道销量比重的数据预测此月份的比重,记为 ξ_i ,权重为 w 。
- 近两个月的 8 次抽查出现不合格状况的均值 μ_i ,权重为 u 。
- 近两个月该渠道出现不合格状况的标准差 β_i ,权重为 v 。

类似种类分析中的方法每一渠道的权重为正态曲线 1 右侧的面积

$$g_i = S_i \quad (27)$$

故每类渠道的抽样比例为

$$T_i = \frac{g_i}{\sum_{j=1}^4 g_j} \quad (28)$$

在实际操作时,先确定出渠道的抽样比例,再在每个渠道中按照不同种类的蔬菜确定其中的抽样比例。

4.3.3 模型求解

利用数据求出的深圳市的模型数据为⁴

表 7: 深圳市蔬菜数据

	豆类	根茎	瓜类	鳞茎类	茄果	叶菜类	芸薹属类
均值	0.81746	0.975914	0.986187	1	0.988095	0.957232	0.931027
标准差	0.071649	0.017322	0.009984	0	0.016836	0.016932	0.051342
销量	0.068523	0.071618	0.162246	0.100796	0.003537	0.419098	0.174182

⁴由于渠道方面新兴渠道的统计数据较少,求方差时不准确,在此仅作常规渠道的分析和对比

设计出深圳市 2017 年 2 月抽样分布

表 8: 深圳市 2017 年 2 月抽样分布

	豆类	根茎	瓜类	鳞茎类	茄果	叶菜类	芸薹属类
抽样比例	0.007626	0.120836	0.277316	0.20691	0.017406	0.04965	0.320256

表 9: 渠道分布

	超市	农贸市场	批发市场
均值	0.960614	0.927998	0.970344
标准差	0.00977	0.028841	0.022981
抽样比	0.320103	0.371421	0.308476

表 10: 模型渠道(由于现有数据不足可能不准确)

	超市	农贸市场	批发市场	点对点配送	网络营销
均值	0.0348162	0.06036	0.033018868	0.0125	0.333333
标准差	0.0097698	0.02884	0.022981497	0.04	0.3
抽样比	0.2710963	0.1911	0.23912797	0.22176	0.076914

五 模型的检验

5.1 抽样模型的检验

由于现阶段没有关于整个市场蔬菜质量的统计,故使用计算机模拟的方法对抽样模型进行检验。

1. 利用上个月的合格率或本月市场的蔬菜质量,把平均质量及其方差通过标准化处理成标准正态分布。
2. 计算检出率,即抽到的不合格样品与不合格商品总数的比
3. 计算样品均值,比较和总体均值的差异

由中心极限定理可知,均值方面的差异具有随机性且不会太大,与蔬菜市场总体的符合度都与抽样的次数呈正相关。

六 进一步讨论

如图为抽样结果检出率的对比,其中蓝色的线条为常规的随机抽样方法,红色的线条为改进后的抽样模型。实现方法为:针对同一个蔬菜质量分布函数,利用本模型的方法(红色线条)和利用原深圳抽样的比例进行 1000 次抽样模拟,求得检出率⁵并进行对比,可见本模型总体上在检出率水平有提升。

⁵检出率 = 抽检到的不合格样品数量/不合格商品总量

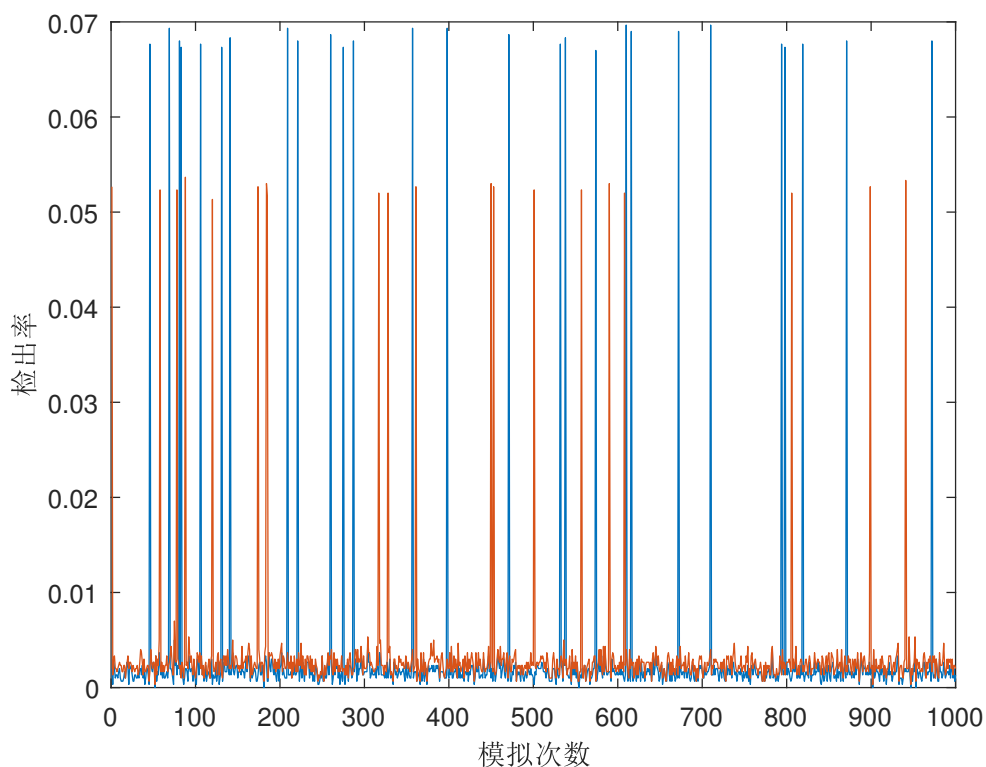


图 3: 抽样检测结果对比

表 11: 深圳市 2017 年 1-3 月抽样计划分布

	罗湖区	大鹏新区	福田区	光明新区	盐田区	龙华新区	南山区	坪山新区	龙岗区	宝安区
农批	0.4	0.64	0.67	0	0	0.29	0.59	0.51	0.41	0.76
农贸	0.20	0.36	0.17	0.61	0.44	0.37	0.24	0.22	0.40	0.19
配送	0	0	0	0	0	0.14	0	0	0	0
超市	0.40	1.00	0.16	0.39	0.55	0.18	0.17	0.27	0.18	0.05

表 12: 深圳市 2017 年 2 月实际抽样比例

	超市	农贸市场	批发市场
实际抽样比	0.472789	0.241497	0.251701

七 模型的评价与推广

本模型以蔬菜为例分析了深圳市蔬菜的安全风险并给出相应的抽检方案, 理论上的表现比较好。也可以类似的推广到其他食品药品抽查与监管。

7.1 模型的不足

1. 由于在实际抽检中往往是以农产品为单位确定抽检场所的比例, 需要对蔬菜, 水果, 肉类的情况进行分析, 本方案具有一定的局限性。

2. 由于每个区的人口不一样, 蔬菜的输入量不一样, 所以需要根据每个区的实际情况建立抽检方案, 将总比例分配到各个区进行抽检。由于没有掌握各个辖区具体的情况, 本模型未能更加详细地把深圳市总体的抽样比例分配到每一个辖区。

7.2 模型的推广

1. 可以将模型类似地推广到食品安全领域, 包括肉制品的监管等
2. 可以将模型推广到其他的地区使用

参考文献

- [1] 周亚. 多属性决策中的 TOPSIS 法研究 [D]. 武汉理工大学,2009.
- [2] 董士县, 论我国食品安全风险评估机制 [J], 山东警察学院学报,2016(3):5
- [3] 杨小敏, 我国食品安全风险评估模式之改革 [J], 浙江学刊,2012(2):4
- [4] 贲智强, 食品安全风险评估的方法与应用 [J], 中国农村卫生事业管理,2012(2):132
- [5] 朱淀, 洪晓娟.2006 2012 年年间中国食品安全风险评估与风险特征研究 [J]. 中国农村观察,2014.2;49
- [6] 李思, 国内外食品安全风险评估机构的比较 [J], 食品工业.2011(10):82
- [7] 刘於, 食品安全综合评价指标体系的层次与灰色分析 [J], 勋河南工业大学,2011(8):77-82
- [8] 苏志海, 张艳, 刘莹, 张硕, 高成忠, 风险矩阵法在出口脱水蔬菜安全风险评估中的应用 [J],2015(2):44-49
- [9] Anna Roccato, Mieke Uyttendaele, Jeanne-Marie Membré: Analysis of domestic refrigerator temperatures and home storage time distributions for shelf-life studies and food safety risk assessment[J], Research International, Volume 96, 2017(6):171-181
- [10] Xiaojia He, Huey-Min Hwang, Nanotechnology in food science: Functionality, applicability, and safety assessment[J], Journal of Food and Drug Analysis, Volume 24, Issue 4, 2016(10):671-681
- [11] 何梓菱, 王哲, 周李, 蔡志杰, 曹沅, 基于动态权值的食品抽检方法 [J], 数学建模及其应用, 2013(9):71-79
- [12] 王三宝, 易秀英, 基于 OC 函数的抽检方案及其风险控制 [J], 2010(7):61-63
- [13] Rounaq Nayak, Patrick Waterson, The Assessment of Food Safety Culture: An investigation of current challenges, barriers and future opportunities within the food industry [J], Food Control, Volume 73, Part B, 2017(3):1114-1123
- [14] 冯杰明, 蔬菜网络销售模式的探索与实践 [J], 中国蔬菜, 2014:10
- [15] 白晨, 卫晓怡, 郭术, 刘斌, 卿蓝, 陆红, 探讨提升流通环节食品抽样检验工作成效的有效机制 [J], 食品工业, 2014(35):6
- [16] 戚建刚, 论基于风险评估的食品安全风险行政调查 [J], 法学家, 2013(5):55
- [17] 周凯, 朱军全, 数学建模竞赛入门与提高 [M], 杭州: 浙江大学出版社, 2011.12:98-150
- [18] 陈杰, MATLAB 宝典 [M], 北京: 电子工业出版社, 2011.11:150-250

附 录

A 数据统计图表

数据搜集:2016 年数据

表 13: 各季节不同种类蔬菜的销量比

	豆类	根茎类	瓜类	鳞茎类	茄果类	叶菜类	芸苔属类
秋季	5.86%	5.86%	17.68%	0.23%	9.53%	43.05%	17.80%
冬季	6.86%	9.50%	14.65%	0.46%	9.84%	42.91%	15.79%
夏季	8.51%	5.61%	16.44%	0.39%	11.41%	41.78%	15.86%
春季	7.08%	6.99%	16.26%	0.36%	10.26%	42.58%	16.48%

表 14: 各季节不同渠道输入比

季节	超市	农贸市场	批发市场	点对点配送	网购	总计
冬季	27.12%	42.10%	26.09%	3.43%	1.26%	100.00%
秋季	31.92%	35.32%	28.13%	3.44%	1.19%	100.00%
夏季	28.05%	35.40%	31.53%	3.87%	1.16%	100.00%
春季	29.03%	37.61%	28.58%	3.58%	1.20%	100.00%

其中不易统计且规模较小的渠道(如流动摊贩等)忽略处理。

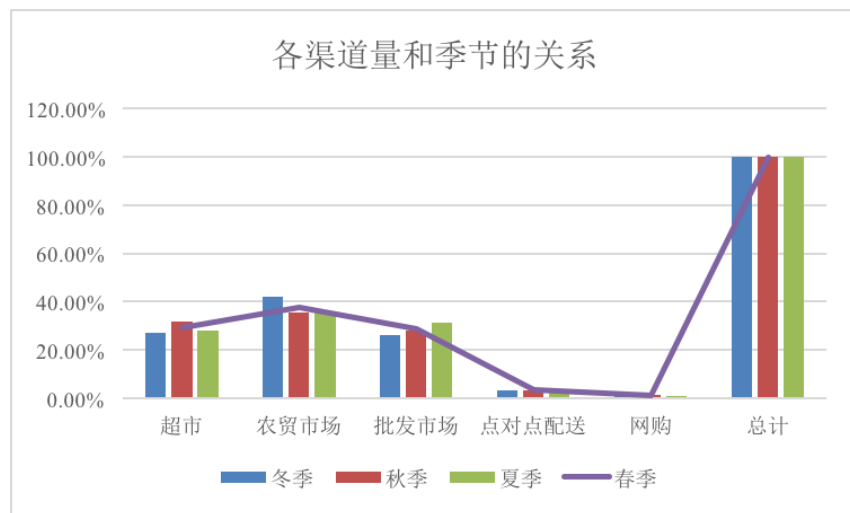


图 4: 渠道数据

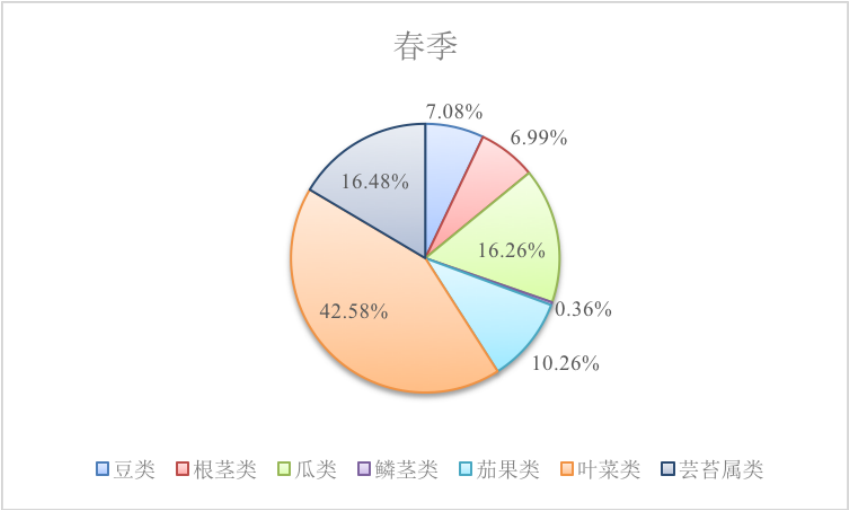


图 5: 春季数据

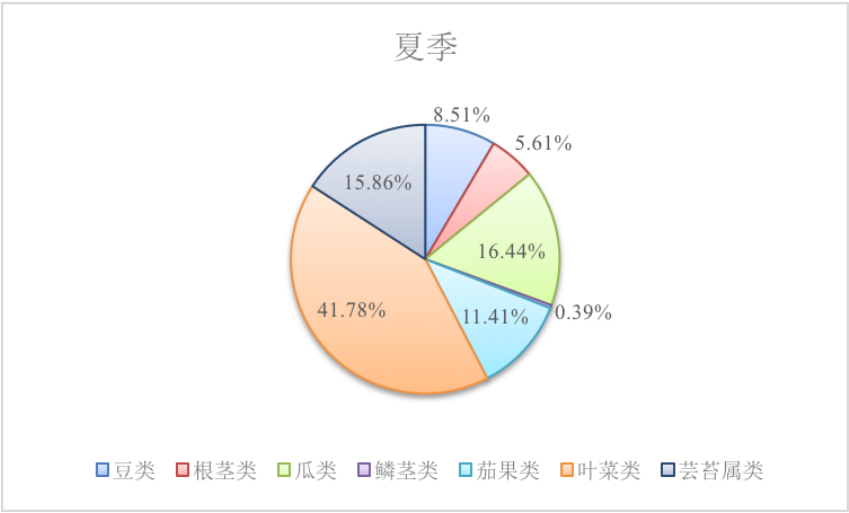


图 6: 夏季数据

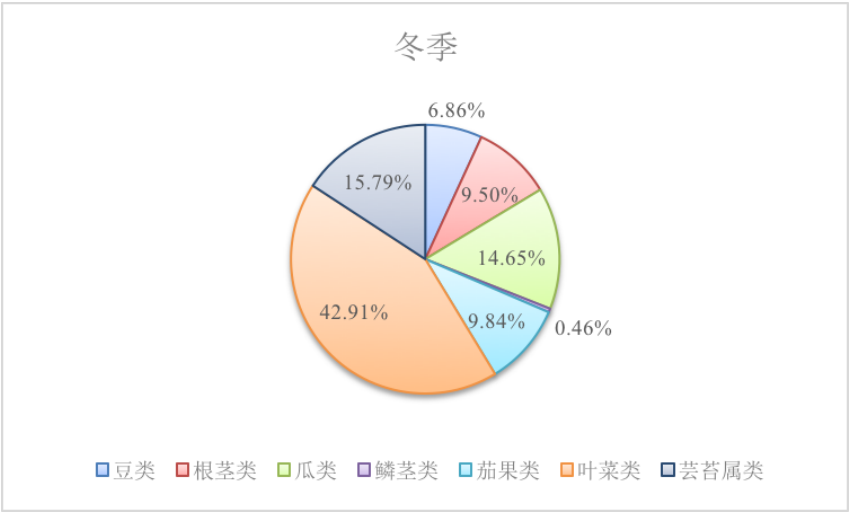


图 8: 冬季数据

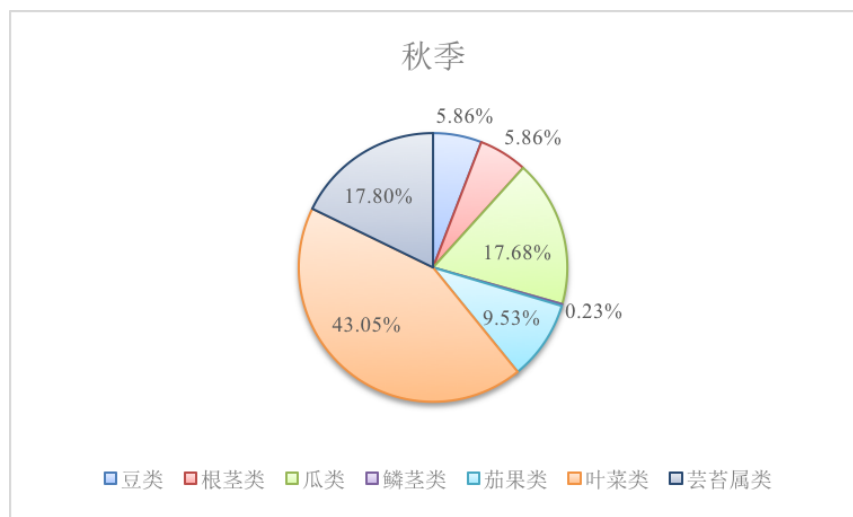


图 7: 秋季数据

B 建议信

深圳市政府各位领导、工作人员：

我们是来自南开大学数学科学院和经济学院的三名大一年级学生。深圳市作为我国改革开放的先驱，几十年来飞速发展的成就和市民生活水平的提高有目共睹，深圳市贸易产业水平在全国也位列前茅。俗语道，民以食为天，食品安全对市民的生活来说更是至关重要，政府相关部门在食品安全风险的监管方面任重道远。然而，在我们对深圳市新鲜蔬菜食品安全相关情况进行调查时，发现了以下几个问题：

- 1. 深圳市新鲜蔬菜输入渠道种类多而杂，且有许多发展迅猛的新兴输入渠道，如互联网销售等；但在有关部门对蔬菜安全风险进行评估时，常常会忽略这些渠道，给蔬菜安全带来一定风险目前的蔬菜安全风险评估体系仍不成熟，深圳市 2017 年第一季度各辖区的抽样比例未实现动态化，多数风险评估工作仍以单纯的少数专家研究为主，在公众之间的普及度较低，某种程度上削弱了风险评估的客观性
- 2. 目前的蔬菜安全风险评估体系仍不成熟，深圳市 2017 年第一季度各辖区的抽样比例未实现动态化，多数风险评估工作仍以单纯的少数专家研究为主，在公众之间的普及度较低，某种程度上削弱了风险评估的客观性。另外，全市各区的抽检工作缺乏统一的执行方案，如《2017 年深圳市食用农产品质量安全监测抽检方案》中提到“各辖区按照农批市场抽样量展总抽样量 65% 以上”就未能完全落实，就 2017 年第一季度而言，深圳市只有福田区和宝安区达到了这个标准。
- 3. 面对蔬菜销售市场复杂多变的现状，现行新鲜蔬菜抽检方案略显程式化，无法适应迅速演变的蔬菜输入渠道；同时，现行抽检方案缺乏针对抽检结果进行动态调整的灵活性，造成抽检结果不能有效警示新鲜蔬菜潜在风险的局面

针对以上问题，我们从建立数学模型出发，结合其他学科的相关知识，建立了评估蔬菜风险的数学模型，设计出了具有一定改进性的新鲜蔬菜抽检方案。基于我们的研究成果，我们向深圳市政府在新鲜蔬菜食品安全监管领域的工作提供以下几点浅显的建议：

- 1. 从蔬菜流量，监察力度，行业素质，流通环节数等角度出发对新鲜蔬菜的输入渠道进行评估，定期监控各渠道的这些指标，其中有指标浮动较大或指标水平明显较低的渠道，对其流经的新鲜

蔬菜进行重点安全监控;同时重点监控互联网销售等新兴输入渠道,可以采取提高考察该渠道各指标水平的频率等方法;在对各渠道进行评估时,从全市新鲜蔬菜的所有输入渠道中选出安全事故最少、风险系数最低的渠道,考察其他渠道与这一最安全渠道的差距,评估其他渠道的安全风险状况

2. 从危险源、市场数量、历史不合格率、渠道风险性等方面,对新鲜蔬菜分类别进行安全风险评估;对于每项宏观的一级指标,采用更细致的微观二级指标进行描述,这些指标的数据或评价来源不能仅限于官方统计的数据,还应该深入市场进行调研后根据市民生活的实际进行汇总,同时提升对新兴食品流入渠道的监控。
3. 在确定抽检方案中各渠道、各种类蔬菜所占比例时,不应局限于既定的方案,而应综合考虑近期抽检结果中各渠道、各种类蔬菜的不合格率、本期抽检时各类蔬菜的预计消费量等动态变化的数据进行应时调整。
4. 在进行蔬菜安全风险评估和抽检方案设计的过程中,有许多数据或实际情况仅靠政府已有的信息和少数专家的研究难以准确把握,因此应该面向社会培养一些食品安全风险评估方面的专业人才,从而获取更多的工作思路和全面客观的数据。
5. 针对目前相关信息的公开在政府部门之间较为孤立,建议把食品安全相关的多个政府部门(如食药监,工商和农业局)的数据结合起来,建立一个统一的数据库,利用大数据分析,及时掌握食品安全情况并对风险进行预测,从而提高深圳市食品安全监管工作的效率,切实保障市民的蔬菜食品安全。

以上建议是我们从学术研究的角度提出的一些简单看法与设想。虽然我们对政府相关工作的开展情况尚不熟悉,以上建议中定然存在一些不够科学可行之处,但是我们仍希望这些研究成果能够对深圳市食品安全监管工作的改进助以绵薄之力。

祝愿深圳市食品安全监管工作开展顺利,不断完善!

南开大学数学科学院 庞敏行 王真
经济学院 王雅晴
2017年8月10日

C 熵权法程序

```
R=[Untitled];  
[rows,cols]=size(R);    % 输入矩阵的大小,rows为对象个数,cols为指标个数  
k=1/log(rows);          % 求k  
  
f=zeros(rows,cols);     % 初始化fij  
sumBycols=sum(R,1);      % 输入矩阵的每一列之和(结果为一个1*cols的行向量)  
% 计算fij  
for i=1:rows  
    for j=1:cols  
        f(i,j)=R(i,j)./sumBycols(1,j);  
    end  
end
```

end

```
lnfij=zeros(rows,cols); % 初始化 lnfij
```

```
% 计算 lnfij
```

```
for i=1:rows
```

```
for j=1:cols
```

```
if f(i,j)==0
```

```
lnfij(i,j)=0;
```

```
else
```

```
lnfij(i,j)=log(f(i,j));
```

```
end
```

```
end
```

```
end
```

```
Hj=-k*(sum(f.*lnfij,1)); % 计算熵值 Hj
```

```
weights=(1-Hj)/(cols-sum(Hj));
```

D 灰色关联分析程序

```
b=[34.2450,0.3862,1.4740,0.8514];
```

```
X=b;
```

```
Y=[116,6,5,1;150,7,6,2;114,8,8,3;14.33,2,6,3;4.81,3,5,1];
```

```
Len=size(Y,2);%取Y矩阵的列数，也就是指标的个数
```

```
Wen=size(Y,1);%取行数，就是目标个数
```

```
%for i=1:Len
```

```
% Y(:,i)=(Y(:,i)-mean(Y(:,i)))/sqrt(var(Y(:,i))); %将Y矩阵用统计方法标准化
```

```
%end
```

```
for i=1:Len-1
```

```
S(:,i)=(Y(:,i)-min(Y(:,i)))/(max(Y(:,i))-min(Y(:,i)));%将Y矩阵标准化，适用
```

```
%越大越好型，把该型指标放在一起，前n-1个，不同情形要修改
```

```
D=(max(Y(:,4))-Y(:,4))/(max(Y(:,4))-min(Y(:,4))); %将Y矩阵标准化，适用于
```

```
%越小越好型，把该型指标放在一起，第n个，不同情形要修改
```

```
end
```

```
SD=[S,D];%把两种不同类型的指标组合在一起
```

```
temp=SD;%给temp变量分配空间，其实可以不分配，只是先分配编译的速度更快
```

```
for i=1:Wen
```

```
temp(i,:)=abs(SD(i,:)-X);%计算评价矩阵与参考序列的差的绝对值
```

```
end
```

```
p=0.5;%分辨系数
```

```
related=Y;%给关联系数 related 变量分配空间
```

```
Min=min(min(temp));
```

```
Max=max(max(temp));
```

```

for i=1:Wen
related(i,:)=(Min+p*Max)./(temp(i,:)+p*Max);
end
f=size(1,Wen);%给关联度分配空间
for i=1:Wen
f(i)=mean(related(i,:));
end
%w=[1/Len 1/Len 1/Len 1/Len 1/Len] %若已知各指标权重，可在此修改
%f=w*related'

```

E 风险模型边界数据

表 15: 危险源(安全部分)

危险源							
	豆类	根茎	瓜类	鳞茎类	茄果	叶菜类	芸薹属类
阿维菌素	0	0	0	0	0	0.190106	0
敌敌畏	0	0.331379	0	0	0	0.145017	0
啉虫脒	0	0	0	0	0	0.337201	0.153144
毒死蜱	0	0.90241	0	0	0	0.548373	0.510476
多菌灵	0.441470925	0	0	0	0	0	0
氟虫腈	0	0	0	0	0	0.452287	0.483124
氟虫腈	0	0	0	0	0	0.533385	0.48976
甲胺磷	0.223579016	0	0	0	0	0.311625	0
甲拌磷	0	0.500865	0	0	0	0	0
甲基异柳磷	0.353785088	0	0	0	0	0.866197	0
克百威	0.541204404	0.252575	0	0	0.194583	0.431519	0.422549
联苯菊酯	0	0	0	0	0	0.270297	0
溴氰菊酯	0	0	0	0	0	0	0.137079
氧乐果	0.719829574	0	0.501153	0	0.385732	0.530349	0.734128
乙酰甲胺磷	0.352931856	0	0	0	0	0	0

表 16: 危险源(危险部分)

	豆类	根茎	瓜类	鳞茎类	茄果	叶菜类	芸薹属类
阿维菌素	0	0	0	0	0	0.760422	0
敌敌畏	0	1.325516	0	0	0	0.580069	0
啶虫脒	0	0	0	0	0	1.348802	0.612577
毒死蜱	0	3.609641	0	0.34	0	2.19349	2.041905
多菌灵	1.765884	0	0	0	0	0	0
氟虫腈	0	0	0	0.09	0	1.809148	1.932497
氟虫晴	0	0	0	0.05	0	2.133539	1.959041
甲胺磷	0.894316	0	0	0	0	1.246499	0
甲拌磷	0	2.003461	0	0	0	0	0
甲基异柳磷	1.41514	0	0	0	0	3.464788	0
克百威	2.164818	1.0103	0	0.553	0.778332	1.726075	1.690196
联苯菊酯	0	0	0	0	0	1.081187	0
溴氰菊酯	0	0	0	0	0	0	0.548316
氧乐果	2.879318	0	2.004612	0.45	1.542929	2.121396	2.936514
乙酰甲胺磷	1.411727	0	0	0	0	0	0