登峰杯

论文类别	☐ 学术作品—自然科学类 ☐ 学术作品—人文社科类 ☑ 数学建模竞赛 ☐ 数据挖掘竞赛 ☐ 艺术创意设计竞赛
论文题目	231100599 外卖固体垃圾对环境影响的多方法评估

清华大学教育研究院
中国高等教育学会学习科学研究分会

摘 要

外卖固体垃圾对环境影响的多方法评估

本文重点探究外卖垃圾与环境、人群特点的关系，并进行估计与定量分析探究。经过数据调研与筛选，并用多种方法对数据进行处理分析并进行建模，最后得到了外卖垃圾与地域，环境及人群的定量关系，最后还对部分建模方法及数据进行了优化处理，进一步提高了数据和结果准确性。

对于第一个问题，采取对固定区域进行问卷调查的获取 148 条数据，再用网上调查 954 数据进行修正，并对数据进行筛选统计。然后求出了每次订外卖平均所需塑料餐盒、塑料袋及纸餐盒数量分别为 2.26，1.67 与 2.14，随后运用区间估计估算出方差并借助 MATLAB 工具计算出了上述数据的置信区间。最后还对数据进行了误差分析：当调查数据有 5%波动时，计算结果会有 10%左右波动。

对于第二个问题，先对国内外固体垃圾处理的现状进行调研，并对目前垃圾处理的三种主要方法填埋、焚烧与堆肥进行了定性分析比较。对于目前国内常用的垃圾处理四种组合方案，根据三种垃圾处理方法在目前所占比例定量分析得出了四种方案单位垃圾所需总成本。然后应用层次分析法给出地下水污染，土壤污染及大气污染三方面的权重，最终求出了填埋、焚烧与堆肥的污染相对指数，根据成本和污染指数最小的原则对四种方案标准化后进行比较，得出转运、焚烧、填埋为当前最佳方案。最后以成本和污染最小为目标，并根据目前垃圾处理情况进行约束，建立多目标规划，应用 LINGO 得出焚烧占比约 41.02%，堆肥占比约 7.96%，填埋占比约 51.02%时为最优解。

对于第三个问题，运用回归分析、判别分析、层次分析模型、支持向量机等方法定量研究外卖产生固体废弃物与居住人群特点之间的关系。几种方法各有优劣，层次分析法容易找到具有何种特质的人产生的废物多。主成分回归基础上的回归得到人群特征到固体垃圾的函数关系。距离判别法根据已有的数据，将的样本在数目上归并，找出各个样本之间公有的信息和特征，并利用贝叶斯公式进行修正。最后综合各个模型结果，得出结论可归纳为：通常女性产生的外卖固体废弃物比男性多，青年人产生的外卖固体废弃物较多，而年老或未成年人产生的外卖垃圾则较少。学历和收入较低的人产生的外卖固体废弃物较多产生的频率更加高，学历和收入较低的人产生的外卖固体废弃物较多，在高学历高收入人群中产生的外卖固体废弃物则较少。

本文进行了误差分析与评估，增加了结论全面与准确性。并在模型改进与优化中探寻出了更加精准与可靠的答案与结论。

本文在前人成果的基础上进行了更加详尽的探究与分析，文中数据处理的严谨性，模型建立的全面，对外卖垃圾问题有一定参考意义。本文不足之处是调查数据随机性不足、数据偏少，对分析结果有一定影响。

关键词：优化 外卖 层次分析 支持向量机 主成分回归 距离判别 贝叶斯判别

1. 前言

1.1 目前研究状况

随着现代化社会的发展，外卖成为了家常便饭。而在为社会带来便捷的同时，外卖所带来的垃圾也对环境造成了污染。外卖类 APP“饿了么”近期发布了中国外卖大数据，显示中国市场用户规模达到 6 亿。据此现状估计，每周最少有 4 亿份外卖飞驰在中国的大街小巷，至少产生 4 亿个一次性打包盒和 4 亿个塑料袋，以及 4 亿份一次性餐具的废弃。数据显示，中国城市生活垃圾堆存量已经超过 80 亿吨。这些垃圾对环境的负面影响包括污染水体、大气、土壤，占用土地，传播疾病等。

固体废弃物处理通常是指通过物理、化学、生物、物化及生化方法把固体废物转化为适于运输、贮存、利用或处置的过程。固体废弃物处理的目标是无害化、减量化、资源化。主要处理技术为焚烧、堆肥、填埋。其它处理技术仅包括热解、气化、水泥窑协同、好氧堆肥、厌氧发酵、湿解等多种方式。虽然这些技术仍未成熟，在工程实践中存在较多问题，但这些项目有助于中国生活垃圾处理技术的创新和探索。

2016 年，中国共有垃圾处理设施 2213 座。从处理技术上分析，卫生填埋和焚烧仍然是中国生活垃圾无害化处理的主要方式，分别占 62%和 30%。其它处理技术仅占垃圾清运量的 2%，包括热解、气化、水泥窑协同、好氧堆肥、厌氧发酵、湿解等多种方式。虽然这些技术仍未成熟，在工程实践中存在较多问题，但这些项目有助于中国生活垃圾处理技术的创新和探索。目前，中国仍有 6%的城市和县城生活垃圾未进行无害化处理，无害化处理能力仍存在 4.5-5 万吨/日的缺口。

1.2 问题重述及研究目的

随着城市生活垃圾量日益增长，固体废弃物处理成为社会关注的热点问题。为对中国城市生活垃圾处理现状进行严谨准确的定量分析，本文通过查阅资料，线上线下问卷调查，了解了外卖数量规模及产生的垃圾量，并对现状进行了分析，提出了更优化可行的方案。

首先，本文在选定的一个特定区域内，根据实地调研的外卖配送数量及所产生的垃圾的数据，估算了各类外卖所带来的固体废弃物的数量。其次，本文根据网上资料数据定量分析各类固体废弃物被处理的情况以及对环境产生的影响，通过对比及优化，提出了更可行的垃圾处理方式。接着，本文结合网上数据以及实地调研的数据，定量研究外卖产生的固体废弃物与选定区域内居住人群的特点之间的联系。最后，本文对中国城市生活垃圾处理现状做了总结， 并提出了优化方案。

2. 模型假设与符号说明

2.1 假设

本文作出了适当的假设以简化问题，但又不会导致失去信息的价值。

1)假设调查得到的所有数据都是真实的,不存在因其他因素而产生的干扰。

2）假设选定的各个区域产生的外卖订单外卖垃圾是稳定的，在一段时期内不会因为任何因素有巨大的波动。

3）假设每一个选项内使用餐盒或塑料袋的人数在选项的区间内是均匀分布的，不存在明显向某一个区段偏移的情况。

4）假设在网上调查得到的数据在地理地域范围上是均一的，不会存在因为地域差别而产生明显的差别。

5）假设外卖使用的如塑料袋，塑料盒，纸餐盒等不存在因制造厂商等因素而产生差别。各类塑料袋，塑料盒，纸餐盒是一样的。

2.2 符号说明

本文使用了较多的符号以清晰和简明文章中的表达。符号说明如表 1，未在表 1 中说明的符号将在文章中进行说明。

表 1 符号说明

符号	定义
$\boldsymbol{x}$	自变量
$\boldsymbol{y}$	因变量
$\boldsymbol{\sigma}$	数据矩阵的方差
$\boldsymbol{\Sigma}$	向量的协方差矩阵
$\boldsymbol{P}$	事件概率
$\boldsymbol{\beta}$	多项式的常数项或者系数
$\boldsymbol{a}$	某一个选项的选择人数
W_A	第二问单位垃圾成本费用
W_B	第二问方案 B 单位垃圾成本费用
W_C	第二问方案 C 单位垃圾成本费用
W_D	第二问方案 D 单位垃圾成本费用
W_t	第二问填埋技术单位垃圾成本费用
W_f	第二问焚烧技术单位垃圾成本费用
W_d	第二问堆肥技术单位垃圾成本费用
r_1	第二问焚烧技术在总垃圾处理技术中的占比
r_2	第二问堆肥技术在总垃圾处理技术中的占比

3. 数据获取

本文所引用的数据主要包括外卖数量、外卖产生的固体废弃物数量、中国城市生活垃圾处理现状、以及不同垃圾处理方式对环境的影响。数据获取的途径主要包括相关机构发布的数据、网上查阅的参考文献资料、以及线上线下问卷调查的第一手数据。问卷部分如下图（完整问卷见附件）：

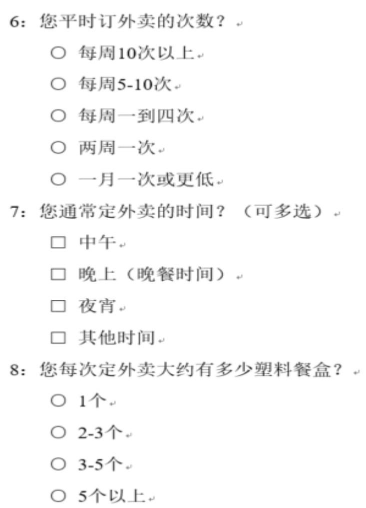

6：您平时订外卖的次数？
○ 每周10次以上
○ 每周5-10次
○ 每周一到四次
○ 两周一次
○ 一月一次或更低
7：您通常定外卖的时间？（可多选）
□ 中午
□ 晚上（晚餐时间）
□ 夜宵
□ 其他时间
8：您每次定外卖大约有多少塑料餐盒？
○ 1个
○ 2-3个
○ 3-5个
○ 5个以上

图 1 问卷（部分）

针对问题一关于外卖所带来的固体废弃物数量的估算和问题三关于分析外卖所带来的固体废弃物与居住人群间的关系，本文以选定区域内实地问卷调查的结果为样本，以线上问卷调查结果为数据库对样本数据分析进行修正。而对于问题二关于固体废弃物的处理现状的定量分析和优化方案，本文主要以相关机构发布的中国城市生活垃圾的处理现状、多种垃圾处理的方式、不同固体废弃物处理方式对环境带来的影响等数据资料作为分析的对象，基于现状提出更优化可行的城市生活垃圾处理方案。

4. 模型建立

4.1 问题一

针对问题一，本文首先考虑特定区域的选取，考虑的因素主要为获取数据的难度，和区域代表性。为降低获取数据的难度，并保证选取数据的精准程度，选定的区域范围不应过大，出入口不应过多。因为较少的出入口可以保证在一段时间内尽可能多的拦截到所要调查的外卖配送人员和订外卖的用户。在区域代表性方面，选定区域应具有一定的外卖流量以避免较大的随机误差，最好是以一种特定人群为主体的区域，如老年人、青年人居住的小区，或是一个写字楼。综合以上因素考虑，本文将调查范围确定在了某公司。经过调查，这一区域中餐有 162 份外卖，晚餐有 53 份外卖。因此中餐塑料盒约有 366-367 个，晚餐约有 120 个。中餐塑料袋约有 260-261 个，晚餐约有 85 个。纸餐盒中餐约有约有 271 个，晚餐约有 89 个。调查得到的数据经过统计后如表 2 所示：

表 2　问卷调查原始数据

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Q1	Q2	Q3	Q4	Q5	Q6_1	Q6_2	Q6_3	Q6_4	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
2	2	2	3	1	3	1	1	1	0	2	2	1	1	1	2	2	1
3	2	2	3	2	3	1	0	0	0	3	2	3	1	1	2	4	1
4	2	2	3	3	3	1	1	0	0	2	1	1	1	1	2	3	1
5	1	2	3	2	2	0	1	0	0	2	2	1	1	1	1	4	1
6	1	3	4	3	3	0	1	0	0	3	3	1	1	1	1	3	1
7	2	2	3	2	3	1	0	0	0	3	1	1	1	1	1	3	3
8	1	2	3	3	3	1	0	0	0	1	1	1	1	1	1	3	3
9	2	3	3	2	3	1	0	0	0	1	1	1	1	1	1	4	1
10	1	3	4	3	3	1	0	0	0	1	1	1	1	1	1	3	3
11	1	3	3	2	3	1	0	0	0	1	1	1	1	1	1	4	3
12	1	3	3	3	3	1	0	0	0	1	1	1	1	1	1	4	3
13	2	2	3	2	3	0	1	0	0	1	1	1	1	1	1	4	3
14	2	1	3	3	2	1	1	0	0	1	1	1	1	1	1	4	3
15	1	3	4	3	3	1	1	0	0	1	1	1	1	1	1	4	3
16	1	0	3	4	3	1	1	0	0	1	1	1	1	1	1	4	3
17	2	0	4	3	3	1	1	0	0	1	1	1	1	1	1	3	3
18	2	0	3	4	3	1	1	1	0	1	1	1	1	1	1	4	3
19	1	0	3	3	3	1	1	0	0	2	1	1	1	1	2	4	3
20	2	3	3	3	3	1	1	1	0	1	1	1	1	1	2	4	3
21	1	3	2	3	3	1	1	0	0	1	1	1	2	2	2	4	3
22	2	2	4	2	3	1	1	0	0	1	1	1	2	2	2	3	3
23	2	3	3	4	4	1	0	0	0	2	1	1	1	1	1	4	1
24	2	2	3	2	2	1	0	0	0	1	2	1	1	1	1	4	1
25	2	4	1	4	2	1	0	0	0	2	2	2	2	1	1	1	1
26	1	1	1	2	4	1	0	0	0	4	3	3	1	1	1	4	1

All_Data_Original

同时，考虑到本文选定区域内数据存在一定误差，因此利用在线问卷调查得到的数据进行了修正。本文将实地调查得到的数据和在线调查得到的数据赋予同等的权重，对数据进行处理，公式如下：

$$A=\frac{n_{在线}}{n_{实地}}\times a_{实地}+a_{在线}$$

其中，A 代表处理过后的这一选项的数据，$a_{在线}$ 代表在线调查获得的某一选项的数目，$a_{实地}$ 代表实地调查获得的某一选项的数目。$n_{在线}$ 代表在线调查获得的参与调查的总人数，$n_{实地}$ 代表实地调查获得的参与调查的总人数。这样处理完后，得到的数据如表 3 所示（表 3 为某一问题的示例，其他数据见附件）：

表 3　问卷处理后数据

Q9	您每次定外	单选题	
	本题选项:		
	1	1-3个	97
	2	4-6个	9
	3	7-9个	3
	4	10个及以上	1

依据此，本文制得各个问题各个选项的统计图如图所示（图 2 为某一问题的示例，其他统计图见附件）：

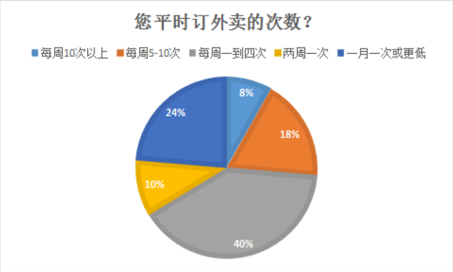


图 2　问题各选项分布图

对于第七问“您每次定外卖大约有多少塑料餐盒”，第八问“您每次定外卖大约有多少塑料袋？”，第九问“您每次定外卖大约有多少纸餐盒？”这三个问题，本文进一步作了点估计和区间估计[1]来对废弃物进行定量的估计。

本文对于每一个选项中餐盒或者塑料袋的数量，均用选项上下限的中值来代表这个选项的每一个人使用的餐盒或者塑料袋的数目。对于没有上界的选项则用选项的下界来代表这个选项每一个人的使用的餐盒或者塑料袋的数目。在第七，八两问中若选择选项 1 则对应的值是 1，若选择选项 2 则对应的值是 2.5，若选择选项 3 则对应的值是 4，若选择选项 4 则对应的值是 5。在第九问中若选择选项 1 则对应的值是 2，若选择选项 2 则对应的值是 5，若选择选项 3 则对应的值是 8，若选择选项 4 则对应的值是 10。

本文首先利用点估计求得每一个点外卖的人使用的餐盒和塑料袋的个数。点估计利用算术平均值公式，即

$$\hat{\mu} = \overline{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

其中，$\hat{\mu}$代表算术平均值的估计值，$\overline{x}$代表求得的这个问题的算术平均值，即点估计的结果。n代表回答这份问卷的总人数，x_i代表第i个人的选项对应的值。求得的结果如下：

表 4　问卷问题七-九的点估计结果

问卷问题	问题内容	结果（个/次）
问题七	每次定外卖的塑料餐盒数量	2.26220794882991
问题八	每次定外卖的塑料袋数量	1.67887576144059
问题九	每次定外卖的纸餐盒数量	2.14449014609801

接着本文求得点估计的方差，公式为：

$$\hat{\sigma}^2 = \sigma^2 = \sum_{i=1}^{n} \frac{(\overline{x} - x_i)^2}{n-1}$$

其中$\hat{\sigma}^2$表示方差的估计值，σ^2表示求得的方差，n代表回答这份问卷的总人数，$\overline{x}$代表上述求得的算术平均值，x_i代表第i个人的选项对应的值。求得结果如下：

表 5　问卷问题七-九的方差

问题七	1.62761856535801
问题八	1.33475670803274
问题九	1.57460238780043

本文在点估计的基础上利用区间估计进一步估算，公式如下：

$$P\{\hat{\theta}_1 < \theta < \hat{\theta}_2\} = 1 - \alpha$$

其中，θ表示已有数据总体的待估参数，P表示概率，$\hat{\theta}_1$表示置信下限，$\hat{\theta}_2$表示置信上限，α表示可信度且满足$0 < \alpha < 1$。因此有

$$P\{\hat{\bar{x}}_1 < \bar{x} < \hat{\bar{x}}_2\} = 1 - \alpha$$

$$P\{{\hat{\sigma}_1}^2 < \sigma^2 < {\hat{\sigma}_2}^2\} = 1 - \alpha$$

利用 MATLAB 程序，本文将可信度α设为 0.95，求得结果如下：

表 6　问卷问题七-九的区间估计结果

问题	平均值$\bar{x}$的置信区间	方差σ^2区间
问题七	[2.26143280891573,2.26298308874409]	[1.6271651117502,1.62826182083831]
问题八	[1.60296275457298,1.60891124875577]	[1.30166457003991,1.30590410543708]
问题九	[1.67820983224883,1.67954169063235]	[1.33436712094773,1.33530931268636]

4.2 问题二

针对问题二，本文首先通过机关机构发布的数据了解了我国城市生活垃圾处理现状。目前垃圾清运量与无害化处理率都在逐步提高，具体趋势如图 3：

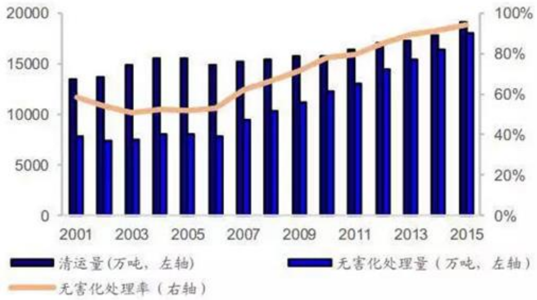


图 3　我国垃圾清运量和处理量趋势

由图 3 可知，2015 年垃圾无害化处理率已达到 90%以上，且仍处于上升趋势。生活垃圾无害化处理方法主要可分为卫生填埋，焚烧发电，及堆肥三种，各自的处理量与占比趋势见图 4：

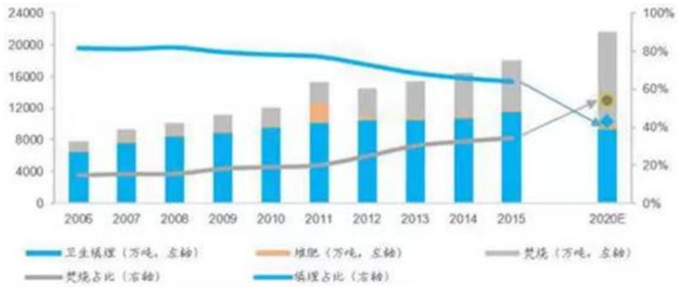


图 4　近年三种主要垃圾处理方式处理量及占比

由图 4 可知，2015 年填埋技术占比约 62%，且处于下降趋势；焚烧技术占比约 30%，且处于上升趋势。三种处理方法各有利弊，各方面的定性比较和相关数据如表 7：

表 7　三种垃圾处理方式的定性比较

处理方式	卫生填埋	焚烧	堆肥

操作安全性	较好，　沼气导排通畅	较好，严格规范操作	较好
选址	较困难，需考虑实地地质条件，防止水体污染，远离市区	选址容易	较困难，因产生恶臭需远离居民区
占地面积	大	小	中
处理工艺	设备简单，操作管理方便，渗滤液处理困难	工艺设备复杂，操作管理要求高，残渣需填埋	管理要求高，处理周期长
产品市场	有沼气回收的填埋场，沼气可发电	热能成电易为本厂，社会利用，经济效益好	推广较难
资源利用	封厂后可恢复土地利用	余热发电，焚烧残渣综合利用	园林绿化
对环境影响	大	最小	较小
投资成本	25—45 元/立方米	35—68 万元/（吨/天）	6—14 万元/（吨/天）
处理成本	25—45 元/吨	50—80 元/吨	40—60 元/吨

由表中数据可知，卫生填埋方法拥有较大的处置能力，但其占用的土地资源较多，对当地的土地污染严重；焚烧处理占用土地资源最少，但建设成本较高，并且产生的大量烟气，过量排放还会对人造成危害。当然还有其他生物处理，化学处理等前沿科技方法处理生活垃圾，但因成本过大，目前普及范围太小以及过往数据缺少，所以在本文中不予以探讨与分析。

本文选取卫生填埋，焚烧，堆肥三种主要的垃圾无害化处理方法，在其所需成本、造成污染两方面，采取层次分析与多目标规划的方法，进行定量分析，最后得出最优化处理方法占比方案。据此，本文首先拟定如下表四种城市生活垃圾处理方式的近似最优方案：

表 8　城市生活垃圾处理方式近似最优方案

方案 A	转运、填埋
方案 B	转运、焚烧、填埋
方案 C	转运、堆肥、填埋
方案 D	转运、焚烧、堆肥、填埋

成本方面，在北京现有的垃圾处理设施中，由于各厂设备技术水平和运行状况存在较大差异。因为应选择平均水平、有代表性的垃圾处理设施项目作为不同垃圾处理方式成本费用比较的依据，并以此作为样本对上文拟定的四种垃圾处理方式的近似最优方案进行分析与评价。结合网上调研的数据，本文选定的代表项目为昌平阿苏卫垃圾焚烧处理厂、南宫垃圾堆肥厂、安定垃圾卫生填埋场。

1 、昌平阿苏卫垃圾焚烧处理厂

该项目由昌昌平区市政管委兴建，位于北京市昌平区小汤山镇阿苏卫，主要解决东城、西城和昌平三区的生活垃圾销纳问题。基本信息如表 9：

表 9　昌平阿苏卫垃圾焚烧处理厂基本信息

总占地面积	65540 平方米
设备占地面积	190 平方米

辅助设施占地面积	750 平方米
日处理垃圾量	240 吨
设计使用寿命	15 年

2、 南宫垃圾堆肥厂

该项目是按照北京市城市总体规划，基本信息如表 10：

表 10 南宫垃圾堆肥厂基本信息

总占地面积	68000 平方米
建筑面积	21600 平方米
厂房面积	15000 平方米
日处理垃圾量	400 吨
年处理垃圾量	13.2 万吨
设计使用寿命	20 年

3、安定垃圾卫生填埋场

该项目是根据北京市城市总体规划。基本信息如表 11：

表 11 安定垃圾卫生填埋场基本信息

总占地面积	21.6 公顷
总填埋容量	356.8 万平方米
设计使用寿命	14 年
日处理垃圾量	700 吨

根据网上资料查阅，由于三个代表项目的规模和占地面积不同，年度总费用不具有可比性，所以本文以单位垃圾成本[2]费用方式计算出三种垃圾处理方式的成本费用作为评判优化垃圾处理方案的评价标准，具体如表 12：

表 12 三种垃圾处理方式成本信息

垃圾处理方式	单位垃圾总成本（元/吨）	单位垃圾运营成本（元/吨）	现今在总垃圾处理中的占比
焚烧	48.46	32.44	30%
堆肥	134.61	76.72	8%
填埋	47.86	27.95	62%

根据上文拟定的四种优化方案，现计算每种方案单位垃圾的总成本费用和运营成本费用，计算公式如下：

$$W_A = W_t$$

$$W_B = r_1 W_f + (1 - r_1) W_t$$

$$W_C = r_2 W_d + (1 - r_2) W_t$$

$$W_D = r_1 W_f + r_2 W_d + (1 - r_1 - r_2) W_t$$

其中，W_A代表 A 单位垃圾成本费用；W_B代表 B 单位垃圾成本费用；W_C代表 C 垃圾成本费用；W_D代表 D 单位垃圾成本费用。

根据上述公式，计算结果如表 13

表 13 四种方案单位垃圾成本费用

方案	单位垃圾总成本（元/吨）	单位垃圾运营

方案 A	47.86	32.44
方案 B	48.04	29.297
方案 C	54.8	31.8516
方案 D	54.98	33.1986

在污染方面，三种处理方法对环境产生的污染主要可分为地下水污染，大气污染以及土壤污染。经过资料搜索与分析提取，本文总结出的卫生填埋，焚烧，堆肥每种方案造成的污染方面及严重程度如表 14：

表 14 三种处理方式对环境产生的污染

处理方式	卫生填埋	焚烧	堆肥
地下水污染	需采取防渗措施，但仍有渗漏可能，污染最大	基本没有	需妥善处理污水，污染较小
大气污染	能通过覆土，导气等措施进行控制，污染较小	可采用先进烟气处理技术达标排放，但成本较高，污染严重	恶臭污染，需除臭设施，污染适中
土壤污染	限于填埋地区，污染较大	基本没有	需控制堆肥中重金属与 PH,污染较小

根据表 14 内容本文对三种方法造成的污染划分严重程度并定义污染等级，污染等级可分为 1 到 5 之间的整数,其中各个等级代表的严重程度如表 15 定义：

表 15　各个等级代表的严重程度

严重程度	轻微	较小	适中	较大	严重
污染等级	1	2	3	4	5

根据对各种处理方案为环境造成影响的定性分析及污染等级的定义，本文将垃圾处理方案造成的污染用等级进行量化，量化结果如表 16 所示：

表 16　三种处理方式对环境的污染的量化结果

处理方法	卫生填埋	焚烧	堆肥
地下水污染	4	1	2
大气污染	2	4	3
土壤污染	4	1	2

将污染量化后，本文将各种污染情况在之后计算中所占比用层次分析法[3]进行赋权和验证，以保证运算的准确性与数据处理的可靠性。本文借助 MATLAB 工具对地下水污染，大气污染与土壤污染在各种方法中造成的污染赋权进行计算。

首先建立层次模型。深入分析实际问题，将有关因素自上而下分层（目标—准则或指标—方案或对象），上层受下层影响，而层内各因素基本上相对独立。其中本文将确定垃圾处理方案定为目标层（O）；将地下水污染，大气污染及土壤污染定为准则层（C）；将卫生填埋，焚烧及堆肥三种方案定为方案层（P）。

然后，用成对比较法和 1~9 尺度，构造各层对上一层每一因素的成对比较阵。本文将分析准则层对目标层的影响并构造准则层的成对比较矩阵。为将成对比较阵量化，本文定义准则层中第i个因素对第j个因素影响程度之比a_{ij}如表 17：

表 17 准则层中第i个因素对第j个因素影响程度之比

尺度	相同	稍强	强	明显强	绝对强

C_i: C_j重要性	1	3	5	7	9

a_{ij}=1，$\frac{1}{3}$，$\frac{1}{5}$，$\frac{1}{7}$，$\frac{1}{9}$时影响程度与上表相反。其中i、j表示准则层中第i、j个因素。本题中共三个影响因素，因此i取值为 1、2 或 3，定义其中 1 代表地下水污染，2 代表大气污染，3 代表土壤污染。特别地，a_{11}，a_{22}，a_{33}默认等于 1。根据数据分析与三种污染的影响关系及上述影响关系定义，本文近似分析得出$a_{12} = \frac{1}{3}$，$a_{13} = \frac{1}{3}$，$a_{23} = 1$，因此可得$a_{21} = 3$，$a_{31} = 3$，$a_{32} = 1$。之后本文用 MATLAB 工具箱构造成对比较矩阵：

$$A = \begin{bmatrix} 1 & \frac{1}{3} & \frac{1}{3} \\ 3 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix}$$

最后，计算权向量并进行一致性检验。对每一成对比较阵计算最大特征根和特征向量，作一致性检验，若通过，则特征向量为权向量。用 MATLAB 工具箱进行进一步分析计算,可求出权向量 Q，即对三种污染的权重，其中地下水污染权重 0.1429，大气及土壤污染权重各为 0.4286。之后计算得一致性指标 CI 与一致性比率 CR 均符合条件。此外，矩阵 A 为一致阵，显然通过一致性检测。

$$\mathrm{CI} = -6.6613\mathrm{e} - 16$$

$$\mathrm{CR} = -1.1485\mathrm{e} - 15$$

对比矩阵 A 通过一致性检验，各向量权重向量 Q 为：

$$Q = 0.1429 \quad 0.4286 \quad 0.4286$$

层次分析过后得到了三种污染的权重，权重之和为 1。因此计算每种垃圾处理方式产生的污染的相对总严重程度可用产生每种污染的污染等级与该污染权重相乘进行求和，算式与结果如下：

$$\text{卫生填埋：} 4 \times 0.1429 + 2 \times 0.4286 + 4 \times 0.4286 = 3.1432$$

$$\text{焚烧：} 1 \times 0.1429 + 4 \times 0.4286 + 1 \times 0.4286 = 2.2859$$

$$\text{堆肥：} 2 \times 0.1429 + 3 \times 0.4286 + 2 \times 0.4286 = 2.4288$$

由此可计算出上文拟定的四种方案的污染值，如下：

方案 A：3.1432

方案 B：2.2859×30%+3.1432×70%=2.88601

方案 C：2.4288×8%+3.1432×92%=3.086048

方案 D：2.2859×30%+2.4288×8%+3.1432×62%=2.828858

综上，从成本和污染两方面去评价四种拟定的方案，结果如表 18：

表 18 四种方案评定结果

方案	成本（元/吨）	污染
方案 A	47.86	3.1432
方案 B	48.04	2.88601
方案 C	54.8	3.086048

方案 D	54.98	2.828858

由表 18 可知，综合成本和污染两方面考虑时，方案 B——转运、焚烧、填埋——为最佳方案。此方案以 2016 年不同垃圾处理方式占比为参考依据，故其评判结果为近似最佳方案。

除上述四种近似最优方案外，本文还将通过上述标化值与 LINGO 多目标规划探寻最优方案。本文现将成本、污染的评定结果标准化，计算公式如下：

$$\partial_i = \frac{W_i}{max\{W_i\}},\ \mathrm{i} = 1,2,3$$

其中，∂表示垃圾处理技术的标准化值，W 表示成本或污染。结果如下：

表 19　成本和污染标准化值

垃圾处理技术	成本标准化值	污染标准化值
焚烧	0.3600	0.7272
堆肥	1	0.7727
填埋	0.3555	1

本文的目标是污染与成本都尽量达到最小，最优化方案的形式将通过每种处理技术占比的方式给出。因为目标是成本与污染都达到最小，所以可将每种技术的标化值相加使其达到最小，再通过若干约束条件求得最优方案。

此处假设在最优方案中焚烧占百分比为x_1，堆肥占百分比为x_2，填埋占百分比为x_3。则目标为：

$$\min 1.0872x_1 + 1.7727x_2 + 1.3555x_3$$

约束条件则有：

$$x_1 + x_2 + x_3 = 100$$

$$x_1，x_2，x_3 > 0$$

此外，根据目前垃圾处理占比趋势，本文还追加了其他约束条件。根据 2016 年处理方式占比数据，填埋处理高于焚烧处理约 30%左右，而填埋占比逐年呈下降趋势，焚烧占比呈上升趋势。因此根据目前处理情况与能力，增加约束条件：

$$x_3 - x_1 > 10$$
$$x_3 - x_1 < 30$$

对于堆肥方法，虽然目前成本较高，尚未普及，但根据近年趋势占比稳定在 5%左右，正在逐步推广。因此增加约束条件：

$$x_2 > 5$$

另外，本文最后目标还需探寻到污染问题单项达到最少情况，以全力解决垃圾环境污染问题，所以本文用 LINGO 程序环境污染标准化数据约束上限进行调试，发现在污染标准化数据总和等于 86 时无解，因此为让污染尽量减小，本文将污染上限定为 87，此时有约束条件：

$$0.7272x_1 + 0.7727x_2 + x_3 < 87$$

此时可得出最优解：

$$x_1 = 41.017609$$
$$x_2 = 7.964785$$
$$x_3 = 51.017609$$

综上，本文提出最优方案为：焚烧占比约 41.02%，堆肥占比约 7.96%，填埋占比约 51.02%。

4.3 问题三

4.3.1 层次分析法[4]

为了在各个因子中进行选择，判断在某个特定性质的人中更可能产生多少垃圾，本文利用层次分析法来实现在复杂的不确定的问题中找到各项权重。本文将人的特质，如年龄，性别，收入，教育程度，所在公司等作为准则层。产生多少垃圾作为目标层，进行仅有一层但有多组的的层次分析。图中表示了使用多少餐盒对年龄的层次分析示意图：

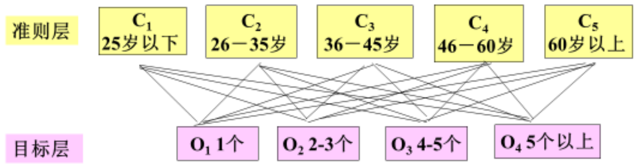


图 5 使用多少餐盒对的层次分析示意图

首先，本文将某一个性质的人产生了多少垃圾的人数，即在某一准则的条件下选择某一目标选项的个数作为这一项的绝对权重，记为w。则依据这个选项本文得到一个权向量$(w_1 \quad \cdots \quad w_n)$($n$为目标层选项的个数)。将这一个性质的人选择产生垃圾相同的人数，$w_i(1 \le i \le n)$互相作比的比值作为成对比较阵的权重，得到成对比较阵A。因为其为一致阵，因此本文不需要进行一致性检验，其必然一致。利用特征根和特征向量公式$Aw = \lambda w$可得其特征向量w，将各个准则的特征向量组合则可以得到同一问题下目标对准则的特征向量阵，及权向量阵。

4.3.2 线性回归法[5]

本文使用的第二种方法是线性回归法。可以将餐盒使用情况作为因变量，各个人的基本信息作为自变量。基于本文调查得到的样本，每一条数据均可以视为由自变量基本信息到因变量餐盒使用情况的映射。由于每一条数据自变量和因变量都是数值，可以根据已有的数据进行学习，通过线性回归得到自变量到各个因变量的函数关系。

本文记性别为自变量x_1，年龄为自变量x_2，教育程度为自变量x_3，年收入为自变量x_4。塑料盒使用情况为因变量y_1，塑料袋使用情况为因变量y_2，纸餐盒使用情况为因变量y_3，餐具使用情况为因变量y_4，餐食剩余情况为因变量y_5，纸巾剩余情况为因变量y_6。自变量和因变量的数值即为各选项的值。

本文利用回归公式

$$y_n = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 (n \in \{n \in N^* | n \le 6\})$$

记自变量矩阵为X，因变量矩阵为Y_n $(n \in \{n \in N^* | n \le 6\})$，系数矩阵为$\beta$，采用最小二乘回归估计方法，公式如下：

$$\beta' = (X^T X)^{-1} X^T Y = \left(\sum x_i x_i^T\right)^{-1} \left(\sum x_i y_i\right) (i \in \{i \in N^* | i \le n\})$$

进行对系数矩阵点估计的求解。通过 MATLAB 程序计算得到的系数矩阵如下：

	塑料盒	塑料袋	纸盒	餐具	餐食	纸巾
β_0	1.378606	1.18038	1.025224	1.001631	1.477334	1.539389
β_1	0.180875	0.055576	0.008579	0.07446	-0.0514	0.065162
β_2	0.033322	0.034247	0.006007	0.057082	0.051542	-0.01299
β_3	0.059538	0.01823	-0.00986	-0.03456	-0.07101	-0.00494

β_4	0.018077	0.002216	0.014431	0.021263	0.011921	0.005794

点估计有一个不足之处是它无法表示所得到的系数的准确度。因此本文利用了区间估计再次进行最小二乘回归。公式如下：

$$P\{\hat{\theta}_1 < \theta < \hat{\theta}_2\} = 1 - \alpha$$

其中，θ表示已有数据总体的待估参数，P表示概率，$\hat{\theta}_1$表示置信下限，$\hat{\theta}_2$表示置信上限，α表示可信度且满足$0 < \alpha < 1$。因此有

$$P\{\hat{\beta}_{n\,1} < \beta < \hat{\beta}_{n\,2}\} = 1 - \alpha \; (n \in \{n \in N^* | n \le 6\})$$

利用 MATLAB 程序，本文将可信度α设为 0.95，求得回归系数上下限如下：

	塑料盒下限	塑料盒上限	塑料袋下限	塑料袋上限	纸盒下限	纸盒上限
β_0	1.132305	1.624907	0.960902	1.399858	0.910536	1.139912
β_1	0.099211	0.262539	-0.01719	0.128347	-0.02945	0.046606
β_2	-0.00504	0.071684	6.27E-05	0.068431	-0.01186	0.02387
β_3	0.011431	0.107645	-0.02464	0.061098	-0.03226	0.012545
β_4	-0.01256	0.048717	-0.02509	0.02952	0.000163	0.028698
	餐具下限	餐具上限	餐食下限	餐食上限	纸巾下限	纸巾上限
β_0	0.84703	1.156233	1.301779	1.652889	1.358263	1.720516
β_1	0.0232	0.12572	-0.10961	0.006809	0.005108	0.125217
β_2	0.033003	0.081162	0.024199	0.078885	-0.0412	0.015225
β_3	-0.06476	-0.00436	-0.1053	-0.03672	-0.04032	0.030438
β_4	0.00203	0.040496	-0.00992	0.03376	-0.01674	0.028327

作出的残差图见附件。当在检查相关系数r^2时，本文发现相关系数如表 20：

表 20 线性回归相关系数

塑料盒	塑料袋	纸盒	餐具	餐食	纸巾
0.027248	0.007408	0.0043	0.032196	0.026184	0.004944

由于相关系数差距过大，因此本文尝试基于主成分回归进行分析。

4.3.3 主成分回归法

本文依然可以像第二种方法一样将餐盒使用情况作为因变量，各个人的基本信息作为自变量。本文试图通过降维，将较多的原始数据信息和变量减少为较少的数据信息和较少变量，而新的变量可以大体上反映原变量的信息。

本文仍然记自变量矩阵为X，因变量矩阵为Y_n $(n \in \{n \in N^* | n \le 6\})$。原变量为$x_p$ $(p \in \{p \in N^* | p \le l\})$，新变量为$z_q$ $(q \in \{q \in N^* | q \le p\})$，记样本的个数为$m$，每个样本的变量数目为$l$，则数据矩阵$X$为：

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1l} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{ml} \end{bmatrix}$$

由于数据的量纲差异较大，范围差异较大，本文需要对数据进行标准化处理。本文采用的是方差标准化法使得标准化后的数据方差为 1，同时进行中心平移使得平均数为 0。公式如下：

$$\overline{x_j} = \sum_{t=1}^{i} \frac{x_{tj}}{i},\quad \sigma_j = \sqrt{\sum_{i=1}^{n} \frac{\left(\overline{x_j}-x_{ij}\right)^2}{n-1}},\quad {x_{ij}}^* = \frac{x_{ij}-\overline{x_j}}{\sigma_j}$$

其中${x_{ij}}^*$表示第i行第j列标准化后的数据，x_{tj}表示第i行第j列标准化前的数据，i表示共有几列，j表示共有几行。

其次，本文建立自变量的相关系数阵R，计算公式和求得的相关系数阵如下：

$$r_{ij} = \frac{\sum_{k=1}^{n}(x_{ki}-\overline{x}_i)(x_{kj}-\overline{x}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ki}-\overline{x}_i)^2 \sum_{k=1}^{n}(x_{kj}-\overline{x}_j)^2}}$$

$$R = \left(r_{ij}\right)_{l\times l}$$

1	0.106472	-0.02209	-0.18079
0.106472	1	0.131604	-0.16217
-0.02209	0.131604	1	0.029231
-0.18079	-0.16217	0.029231	1

然后，本文求矩阵R的特征根$\lambda_q\ (q \in \{q \in N^* | q \le l\})$（对$\forall 1 \le x < y \le q$都满足$\lambda_x > \lambda_y$）和特征向量$a_q\ (q \in \{q \in N^* | q \le l\})$，以确定与相关矩阵的$q$个较大的特征值所对应的特征向量相等的原变量$x_p$在各个主成分新变量$z_q$上的荷载$a_{pq}$。$a_{pq}$是特征向量$a_q$的第$p$个值。公式如下：

$$RA = \lambda A$$

其中A表示各个特征向量，λ表示各个特征根，求得的特征根矩阵如下：

0.773125	0.840018	1.079256	1.307601

特征向量矩阵如下：（每列为一个特征向量）

-0.17622	0.755247	-0.32618	-0.54051
-0.61883	-0.38611	0.381381	-0.5679
0.423931	0.349138	0.822693	-0.14684
-0.6374	0.398274	0.267078	0.603136

利用贡献率公式和累计贡献率公式

$$\frac{\lambda_i}{\sum_{k=1}^{q}\lambda_k}\ (i=1, 2, \dots, p) \quad 和 \quad \frac{\sum_{k=1}^{i}\lambda_k}{\sum_{k=1}^{q}\lambda_k}\ (i=1, 2, \dots, p)$$

求得直到第三主成份累计贡献率为0.806719，大于0.8，因此取前三个特征值作为主成分。设主成分为：

$$z_1 = a_{11}x_1 + a_{21}x_2 + a_{31}x_3 + a_{41}x_4$$

$$z_2 = a_{12}x_1 + a_{22}x_2 + a_{32}x_3 + a_{42}x_4$$
$$z_3 = a_{13}x_1 + a_{23}x_2 + a_{33}x_3 + a_{43}x_4$$

其中a为

-0.54051	-0.5679	-0.14684	0.603136
-0.32618	0.381381	0.822693	0.267078
0.755247	-0.38611	0.349138	0.398274

利用上述最小二乘回归，点估计和区间估计方法，由回归公式

$$y_n^* = \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 (n \in \{n \in N^* | n \le 6\})$$

得到系数矩阵β如下：

	塑料盒	塑料袋	纸盒	餐具	餐食	纸巾
β_1	-0.20454	-0.16981	-0.1595	-0.22205	-0.17519	-0.17574
β_2	0.268683	0.244276	0.354158	0.269743	0.315164	0.291636
β_3	0.400675	0.292137	0.467573	0.369267	0.323962	0.476837

	塑料盒下限	塑料盒上限	塑料袋下限	塑料袋上限	纸盒下限	纸盒上限
β_1	-0.25638	-0.1527	-0.22186	-0.11776	-0.21677	-0.10224
β_2	0.220762	0.316604	0.19616	0.292392	0.301224	0.407092
β_3	0.350522	0.450828	0.241779	0.342494	0.412173	0.522973
	餐具下限	餐具上限	餐食下限	餐食上限	纸巾下限	纸巾上限
β_1	-0.27623	-0.16787	-0.23183	-0.11854	-0.2317	-0.11978
β_2	0.219655	0.31983	0.262799	0.367528	0.239903	0.34337
β_3	0.316847	0.421688	0.269158	0.378765	0.422693	0.53098

最后进行逆标准化变换后得到的用原变量表示的回归方程

$$y_n = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 (n \in \{n \in N^* | n \le 6\})$$

得到系数矩阵：

	塑料盒下限	塑料盒上限	塑料袋下限	塑料袋上限	纸盒下限	纸盒上限
β_0	1.992579	1.992579	1.423933	1.423933	1.069573	1.069573
β_1	0.448787	0.433149	0.285051	0.271199	0.20587	0.19792
β_2	0.059731	0.02112	0.059945	0.025746	0.022947	0.003317
β_3	0.275291	0.354778	0.197871	0.268274	0.157076	0.197486
β_4	0.022146	0.086706	0.006613	0.063796	0.026417	0.059239
	餐具下限	餐具上限	餐食下限	餐食上限	纸巾下限	纸巾上限
β_0	1.231911	1.231911	1.35436	1.35436	1.607607	1.607607
β_1	0.270195	0.259909	0.234369	0.222196	0.360714	0.348438
β_2	0.047083	0.021686	0.057646	0.02759	0.027531	-0.00278
β_3	0.168294	0.220576	0.197587	0.259462	0.222029	0.284422
β_4	0.005789	0.048253	0.013488	0.063743	0.033965	0.084641

4.3.4 距离判别法[6]

本文尝试依据已有的数据，随机抽样$\frac{2}{3}$作为学习样本，$\frac{1}{3}$作为测试集，将庞大的样本在数目上高度归并，找出各个样本之间公有的信息和特征，从而得出类似的废弃物情况时人群的共同特征从而确定关系。

由于本文调查得到的数据在自变量相同时对应的因变量却不同，但判别时需要各个样本不同，因此本文将相同自变量时的因变量取代数平均值，作为自变量时因变量的代表。对于塑料盒，塑料袋，纸盒使用数量这三个问题将选项的值转化为问题的数量，如有几个塑料盒，进行求数量平均值。对于餐具，餐食，纸巾这三个问题有剩余则赋 0，无剩余则赋 1. 当不存在这样的自变量组合时因变量补 0。由于取得的代数平均值并不为整数，因此本文将求得的代数平均值对更加接近的选项的值进行四舍五入，再次转化为原始选项的值。转化出来的选项如表所示。然后本文将原始自变量组合对应的因变量的值用转化后的值替代，用以接下来进一步的分析。如表所示。

用马氏距离判别法，公式如下：

$$d(x,y)=\sqrt{(x-y)\Sigma^{-1}(x-y)^T}$$

其中x，y均表示两个行向量，Σ表示协方差矩阵，$d(x,y)$表示求得的数据的马氏距离。。

本文随机抽取的判别后的结果与最初的调查结果比对如下表。判别准确率如下表。由于各个问题的判别准确率均较低，不足以准确的反应各个变量之间的特征，因此本文考虑采用另外的方法。

塑料盒	塑料袋	纸餐盒	餐具	餐食	纸巾
0.385042	0.411111	0.5528	0.7167	0.6	0.5583

4.3.5 贝叶斯判别法[7]

在以上的距离判别方法中，没有考虑到各个样本在总体中出现的频率多少，也没有考虑到误判带来的损失。而贝叶斯判别法在距离判别的基础上进行了修正，公式如下：

$$P(B_i\mid A)=\frac{P(A\mid B_i)P(B_i)}{\Sigma P(A\mid B_i)P(B_i)}$$

其中$P(B_i|A)$表示后验概率，$P(A|B_i)$表示先验概率，$P(B_i)$表示样本出现的频率，Σ 表示总体的协方差矩阵。判别规则有后验概率最大原则和平均错判损失最小，判别规则如下：若满足：

$$P(G_l\mid x_0)=\frac{p_l f_l(x_0)}{\Sigma p_j f_j(x_0)}=\max_{1\le i\le k}\frac{p_i f_i(x_0)}{\Sigma p_j f_j(x_0)}$$

则将x_0判给G_l，其中G_i为总体，$f(x)$为总体G_i的概率密度函数，p_i为G_i的先验概率，即样本x_0发生时它属于某类的概率。k为总体G_i的个数。判别分析的解公式如下：

$$ECM=\sum_{i=1}^{k}p_i\sum_{j\neq i}C(j/i)P(j/i) \qquad p(j/i)=P(X\in D_j/G_i)=\int_{D_j}f_i(x)dx \qquad i\neq j$$

其中用$P\left({}^{j}/_{i}\right)$表示将来自总体$G_i$的样品错判到总体$G_j$的条件概率，$C\left({}^{j}/_{i}\right)$为这种错判所造成的损失。$D_k$为待判样本的一个分划。$ECM$是平均错判损失。使$ECM$最小的一组解是贝叶斯判别分析的解。

利用 MATLAB 程序，本文依然随机抽样$\frac{2}{3}$作为学习样本，$\frac{1}{3}$作为测试集，进行贝叶斯判别求解。由于按照上一节所述方法转换过后的数据不存在选项 4，因此选项 4 的条件概率为 0。计算后发现除了纸盒的问题存在多种判别结果外，

其余 5 个问题仅存在一种判别结果，不能从中分类归纳得出结论，因此本文将原始数据与判别结果相同的去掉，再次进行判别分析。

在剩余餐纸的情况当中，表中显示了判别结果为 1，也就是有剩余餐纸扔掉的选项，与人群信息的关系。依据表中信息可以得出男性较女性更通常将剩余的餐纸扔掉。可以得出年龄低的人群更易于将餐纸扔掉而成为外卖垃圾。可以得出学历较低的人群更加会把餐纸扔掉。可以得出尚未工作和高收入人群更倾向于扔掉餐纸。

性别	判别结果分类	总计	学历	判别结果分类	总计
男	102	434	高中及以下	31	61
女	24	519	大专	5	77
年龄	判别结果分类	总计	本科	55	463
25 岁以下	88	128	研究生及以上	35	352
26—35 岁	37	222	年收入	判别结果分类	总计
36—45 岁	1	330	0-1 万元	0	70
46—60 岁	0	236	1 万元-10 万元	0	259
60 岁以上	0	37	10 万元到 20 万元	12	311
			20 万元到 35 万元	21	129
			35 万元以上	18	100
			还没工作	75	84

在进行第二次贝叶斯判别以后，发现对于第九问由于选项过于极重，去掉后样本量过小，因此无法进行分析。对于第十，第十一两问因为选项只有两个，去掉后选项划分，即需要判别的类别仅剩一类，因此无法进一步进行判别分析。对于有多个选项的第七、八个问题第七问获得了 0.7766 的回判准确率，第八问获得了 0.9405 的回判准确率。第二次贝叶斯判别结果如表所示。将两次判别的结果综合如表所示，第一个表是第七问选项为 1，也就是经过转换后通常每次外卖使用 1 个塑料盒的人的特征。第二个表是第七问选项为 3，也就是经过转换后通常每次外卖使用 4-5 个塑料盒的人的特征。第三个表是第八问选项为 2，也就是经过转换后通常每次外卖使用 2-3 个塑料袋的人的特征。

性别	判别结果分类	总计	比率	学历	判别结果分类	总计	比率
男	57	434	0.131336	高中及以下	3	61	0.04918
女	15	519	0.028902	大专	12	77	0.155844
年龄	判别结果分类	总计	比率	本科	15	463	0.032397
25 岁以下	9	128	0.070313	研究生及以上	41	352	0.116477
26—35 岁	22	222	0.099099	年收入	判别结果分类	总计	比率
36—45 岁	27	330	0.081818	0-1 万元	8	70	0.114286
46—60 岁	8	236	0.033898	1 万元-10 万元	7	259	0.027027
60 岁以上	6	37	0.162162	10 万元到 20 万元	10	311	0.032154
				20 万元到 35 万元	10	129	0.077519
				35 万元以上	25	100	0.25
				还没工作	12	84	0.142857

分析如下：在使用塑料盒较少的选项中，男性占比远多于女性。在老年人中使用的餐盒较少。同时在学历较高的人中占比较高，使用的餐盒较少。在尚未工作或者年收入高的人群中使用的餐盒较少。

性别	判别结果分类	总计	比率	学历	判别结果分类	总计	比率
男	7	434	0.016129	高中及以下	4	61	0.065574
女	15	519	0.028902	大专	11	77	0.142857
年龄	判别结果分类	总计	比率	本科	7	463	0.015119
25 岁以下	9	128	0.070313	研究生及以上	0	352	0
26—35 岁	1	222	0.004505	年收入	判别结果分类	总计	比率
36—45 岁	2	330	0.006061	0-1 万元	8	70	0.114286
46—60 岁	3	236	0.012712	1 万元-10 万元	11	259	0.042471
60 岁以上	7	37	0.189189	10 万元到 20 万元	3	311	0.009646
				20 万元到 35 万元	0	129	0
				35 万元以上	0	100	0
				还没工作	0	84	0

在使用塑料盒较多的选项中，女性占比远高于男性，年龄大的人使用的较多，学历和收入较低的人占比较高，而学历高或者收入高的人的人则极少使用大量餐盒。

性别	判别结果分类	总计	比率	学历	判别结果分类	总计	比率
男	56	434	0.129032	高中及以下	6	61	0.098361
女	103	519	0.198459	大专	34	77	0.441558
年龄	判别结果分类	总计	比率	本科	73	463	0.157667
25 岁以下	11	128	0.085938	研究生及以上	46	352	0.130682
26—35 岁	21	222	0.094595	年收入	判别结果分类	总计	比率
36—45 岁	63	330	0.190909	0-1 万元	16	70	0.228571
46—60 岁	56	236	0.237288	1 万元-10 万元	34	259	0.131274
60 岁以上	8	37	0.216216	10 万元到 20 万元	52	311	0.167203
				20 万元到 35 万元	45	129	0.348837
				35 万元以上	10	100	0.1
				还没工作	2	84	0.02381

在使用塑料盒较多的选项中，女性占比远高于男性，年龄大的人使用的较多，学历和收入较低的人占比较高，而学历高或者收入高的人的人则较少使用大量塑料袋。

综上所述，综合各个结果可以发现，通常女性产生的外卖固体废弃物比男性多，年龄在青年人中产生的外卖固体废弃物较多，而年老或者未成年人的人则很少。学历和收入较低的人产生的外卖固体废弃物较多产生的频率更加高，因此学历和收入较低的人产生的外卖固体废弃物较多，在高学历高收入人群中产生的外卖固体废弃物则较少。

4.3.6 支持向量机

支持向量机（SVM）是一种在解决非线性问题中有极大优势的算法，其将低维数据映射到高维空间中并找到最优超平面作为将给定数据分类的标准。

本文使用 MATLAB 和 LIBSVM-3.22 [8]进行了支持向量机的计算，其示意图如下：

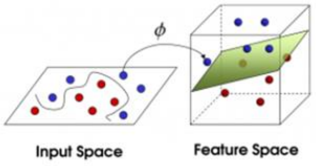


图 6 支持向量机示意图

本文从所有数据中抽取$\frac{1}{3}$作为测试数据，$\frac{2}{3}$作为训练数据训练出一个 SVM 模型并通过计算计算已知样本点的错判率和在训练集上的错误率而评估出模型的切合程度。

使用的 MATLAB 代码见附录。最后计算出模型在训练集上的准确度为~63%，在测试集上的错误率 err_rate 是38.02%。此处列举了模型认为分类正确概率最大的 10 组数据：

表 21 模型认为分类正确概率最大的 10 组数据

第一问	第二问	第三问	第四问	第五问-预测	是否正确
1	1	1	2	2	是
2	1	2	3	3	是
2	2	4	5	2	部分正确（多解）
2	3	2	6	3	是
1	3	4	6	2	否
2	2	2	4	2	是
2	4	3	1	2	部分正确（多解）
2	3	3	5	2	部分正确（多解）
2	3	4	5	3	正确
2	4	2	1	2	部分正确（多解）

由于有时会出现前四题作答完全一样但是第五题选择内容不一样的数据，因此这时模型只能给出一个可能的选择；同样，这些数据也干扰了模型的学习，因而无法取得大于64%正确率的 SVM 模型，这也是本方法的局限所在。

最终观察各组数据，得出结论：训练模型显示，拥有大专学历的 26 至 45 岁之间的女性在外卖上产生的垃圾比拥有其他学历或其它年龄段女性产生的更多。

4.4 问题四

“亲，您的外卖已送达。”

又是一个雾霾天儿，外卖小哥将你的订单“风雨无阻”送到门口。你懒洋洋的从沙发上起来接过一塑料袋的餐盒——烤串儿，小龙虾....... 等等！先别急着吃。

数一数这一餐又将产生多少外卖垃圾呢？据本文统计，全国约有**60%以上**的人每周至少订一次外卖，而每次用餐就平均会**消耗 2.26 个塑料餐盒，1.68 个塑料袋和 2.14 个纸餐盒**！你可曾想过，每周有多少外卖包装，载着热腾腾的美食到你家做客，却又甘心被你丢弃？你可曾注意过，他们消失在垃圾桶里的身影，又将何去何从？

目前国内生活垃圾无害化处理方式主要分为三类：卫生填埋，焚烧与堆肥，其中前两种总和占比可达 90%以上，堆肥因成本较高尚未大规模普及，但不论哪种处理方式，都会对地下水，大气和土壤造成一定污染。**二噁英过量排放，重金属污染土壤，甚至是垃圾恶臭，这些都在严重影响着我们人体的健康。**

你穿的带有弹性，雪纺，化纤面料的衣服，在经过水洗之后，它在水中的塑料微粒经过水循环系统流入江河湖海。
↓
这些塑料微粒会吸收海洋内的有毒污染物（例如不易降解的多氯联二苯等）。
↓
又由于它们的形态和鱼卵极为相似，易被其他鱼类误食。
↓
最终海鱼又经过捕捞回到人类的餐桌，顺利实现食物链系统循环。

有些塑料污染，甚至还会通过食物链，重新回到我们的餐桌上！

是不是吓得你不敢再订外卖了？别害怕。减少外卖垃圾需要你我的共同努力，关键并不在少订外卖，而是需养成更好的用餐习惯。

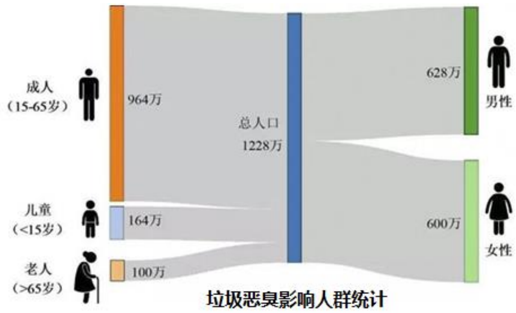
据本文统计分析，女士订外卖所用餐盒数量普遍多于男性，而高学历人士所用餐盒数量也相对更低。是不是开始感叹要保护环境还得先提高自己知识水平了？

当然，最靠谱的还是要从身边小事入手，如在家定外卖时，是不是应该注意一下那个“零餐具”选项，而改用自己家中的餐具呢？

除此之外，在吃完外卖后，也不要一股脑都扔到垃圾桶中。时刻提醒自己注意外卖垃圾的分类与回收，分清一次性餐具与可回收垃圾，从源头起减少废弃物数量。

减少外卖垃圾还更需要多方努力，如外卖公司可以致力研发环保型餐具餐盒，减少一次性外卖垃圾的产生。同时还应减少外卖餐盒的配送量，做到能少装时就少装，这样同样可以达到很好的效果。

可降解餐盒也是一种保护环境的思路，但其降解过程也需要一定周期，同时制造中可能也会产生一定污染。环保人士认为，环境友好性还应从全生命周期角度审视，而这就包含了外卖物流，废弃物排放，处理处置等多项环节。多方在共同解决垃圾围城问题时，也应从这些角度进行更深入的思考。



总之，减少外卖垃圾，保护环境的责任需要我们每个人共同承担。改变自己的习惯，从点滴做起，你就是在保护我们共同的家园。

5. 灵敏度分析

对第一问进行误差分析时，本文将使用的外卖包装数目扩大 5%，订单量也增加 5%，或将使用的外卖包装数目减少 5%，订单量也下降 5%。得到上升时的餐盒变化率大约为 10.25%。中餐约使用 404-405 个塑料盒，晚餐约使用 132-133 个塑料盒。中餐塑料袋约有 287-288 个，晚餐约有 94 个。纸餐盒中餐约有约有 299 个，晚餐约有 98 个。下降时餐盒变化率大约为 9.75%。中餐约使用

330-31 个塑料盒，晚餐约使用 109 个塑料盒。中餐塑料袋约有 236-237 个，晚餐约有 94 个。纸餐盒中餐约有约有 246 个，晚餐约有 88 个。

在对层次分析进行灵敏度分析时，由于使用的外卖包装和各个选项均为整数，因此在取整后最终结果不变，得出模型比较稳定。在对贝叶斯判别进行误差分析时，对多数问将数据上浮或下浮 5%时最终四舍五入得到的结果是相同的，因此这个模型比较稳定。

对于第二问，误差分析如下：

VARIABLE	CURRENT COEF	ALLOWABLE INCREASE	ALLOWABLE DECREASE
X1	1.087200	0.769013	INFINITY
X2	1.772700	INFINITY	0.551350
X3	1.355500	1.102700	INFINITY
X1,X2,X3	0.000000	INFINITY	0.000000

经过 LINGO 程序进行的灵敏度分析，本文得出 x1，x3 在约束条件下可增加空间分别 0.769013 和 1.102700，x2 的可减少空间为 0.551350。由此可见在污染总指标约束尽量小的不断调试中，解的变动范围已经大幅缩小，进一步说明得到结果的优化性与准确性，为第二问最后结论的准确得出做下铺垫与基础。

6. 总结

在问题二的层次分析模型中,明显优点为通过建模方法可以让结论获得数据与分析支撑,因为污染不容易进行量化,而通过过层次分析的方法可以将各类污染进行量化计算。对每种污染进行权重赋值,可以让分析结果更加具有科学性。但不足之处是在污染等级量化时是在查阅资料并对比分析后进行的人为污染等级赋值，而非完全依靠数据支撑可能缺少客观性,但碍于各个标准无法进行统一，污染量化中的人为赋值也无法完全避免。

对于问题三，优点为判别模型稳定，不会因为少数数据的不准确而影响正确率；因为使用了将低维数据映射到高维的方法，而获得了较好的对非线性数据的判别能力，正确率较高。缺点则是由于本方法的缺陷，在模型的训练过程中，无法避免数据源中一对多的数据导致的误差，因而正确率无法突破 70%。

7. 参考文献

[1] 用 MATLAB 进行区间估计与线性回归分析
[2] 北京城市生活垃圾处理策略研究_刘凯
[3] MATLAB 求解层次分析法程序代码
[4] 基于 AHP 法对外卖网站的综合评价——以“饿了么”网站为例
[5] MATLAB 主成分分析
[6] Fisher 判别和 Mahalanobis 距离判别比较研究
[7] 判别分析和 SPSS 的使用
[8] Chih-Chung Chang and Chih-Jen Lin, *LIBSVM : a library for support vector machines.* ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at *http://www.csie.ntu.edu.tw/~cjlin/libsvm*

8. 附录

调查问卷的副本

1：您的性别

○ 男 ○ 女

2：请问您的年龄是？

○25岁以下 ○46－60岁

○26－35岁 ○60岁以上

○36－45岁

3：受教育程度

○ 高中及以下 ○ 本科

○ 大专 ○ 研究生及以上

4：请问您的收入？

○ 0-1万元 ○ 20万元到35万元

○ 1万元-10万元 ○ 35万元以上

○ 10万元到20万元

5：您平时订外卖的次数？

○ 每周10次以上 ○ 两周一次

○ 每周5-10次 ○ 一月一次或更低

○ 每周一到四次

6：您通常定外卖的时间？（可多选）

□ 中午 □ 夜宵

□ 晚上（晚餐时间） □ 其他时间

7：您每次定外卖大约有多少塑料餐盒？

○ 1个 ○ 3-5个

○ 2-3个 ○ 5个以上

8：您每次定外卖大约有多少塑料袋？

○ 1个 ○ 3-5个

○ 2-3个 ○ 5个以上

9：您每次定外卖大约有多少纸餐盒？

○ 1-3个　　○ 7-9个

○ 4-6个　　○ 10个及以上

10：您订外卖大多使用一次性餐具还是自己的餐具？

○ 一次性餐具　　○ 自己的餐具

11：您订外卖通常有剩余餐食扔掉吗？

○ 通常扔掉　　○ 留着继续吃或者当餐吃完

12：您订外卖通常有纸巾扔掉吗？

○ 通常扔掉　　○ 留着继续用或者没有剩余

13：处理外卖垃圾时您有将垃圾分类的习惯吗？

○ 经常　　○ 很少

○ 偶尔　　○ 从来没有

14：处理外卖垃圾时您如何处理垃圾？

○ 一起扔掉，因为没有分类垃圾桶

○ 分类回收

○ 有垃圾桶但是也一起扔掉

官方网站	www.dengfengbei.com
微信公众号	**Dengfengbeijingsai**
官方 QQ 群	(1) “登峰杯”学术作品学生 QQ 群 571526693 (2) “登峰杯”数学建模学生 QQ 群 571535826 (3) “登峰杯”机器人学生 QQ 群 571540979 (4) “登峰杯”结构设计学生 QQ 群 592858677 (5) “登峰杯”数据挖掘学生 QQ 群 144821810 (6) “登峰杯”艺术创意设计学生 QQ 群 318850726
官方邮箱	dengfengbei@126.com
联系电话	010-52909593，18310079788 (工作日 9:00~12:00，13:00~17:00)