长安大学2019年研究生数学建模竞赛

控 制 页

选择的题号（从A/B/C/D中选择一项填写）：D

参赛队成员（打印并签名，注明专业）：

1 曹华科 班级：2018126002

2 程旺盛 班级：2018126004

3 马劼 班级：2018126006

队长：曹华科

联系电话：18802975700

邮箱：496405242@qq.com

日期： 2019年 5 月 22 日

**摘要**

本文针对电视台的广告推送营运问题和产品销售或制造商的收益问题，运用多种数学建模方法建立了基础问题数学模型，并且运用MATLAB等软件进行编程，最终根据建立的竞价交易模型对编撰的数据进行试算，得出合理结果。

针对问题一，要求依据视频广告和电视频道用户的分类特征建立二者分类匹配推送静态模型。我们通过提取电视频道用户的年龄特征和视频广告的多种类型利用层次分析法建立层次结构模型，运用了MATLAB编程计算了判断矩阵并进行一致性检验，最终得出不同的广告对于不同年龄段的用户的匹配度。我们得出的结论符合客观情况。

针对问题二，要求设计卖方分时段竞卖合理底价估算模型。处理时分析得到现有主流媒体的收视率和底价之间的关系，加以考虑到卖方的主要分类和影响因素的不同：通过Fortran建立与时段活跃度、电视台活跃指数、电视节目收视人群及其经济实力有关的每个电视台的收视率底价模型，利用分析得到的各影响指数的大小，最终可以通过人机交互得到预估底价。

针对问题三，要求基于已知电视频道用户收视历史情况和在播视频广告的产品销售情况，并进一步对用户和广告进行特征提取，再根据相应的特征建立起一个大型的矩阵，将该矩阵输入到BP神经网络中进行训练，从而得到一个对建立视频广告与频道用户的分类匹配推送更新模型。

针对问题四，要求建立竞价交易模型，并在极大化卖方收益的同时提升收视率和买方产品销售量，我们首先采用了多元线性回归方法，其次采用了主因分析法，进一步构建了对竞拍价预测的BP神经网络，最后采用差分进化算法和神经网络算法相结合，当输入自变量电视台的底价以及宣传费用时，可以得出电视台的最大收益。

针对问题五，我们通过上述已经建好的模型，对收集到的实际数据进行求解，并得到符合客观情况且可信度高的结果。

目录

[一、问题的重述 1](#_Toc11637)

[二、问题分析 1](#_Toc10334)

[三、模型假设 2](#_Toc24964)

[四、符号约定 2](#_Toc11106)

[五、问题一的解 4](#_Toc21768)

[5.1 问题分析 4](#_Toc19653)

[5.2 层次分析法构造模型 5](#_Toc10287)

[5.3 模型求解 9](#_Toc11269)

[六、问题二的解 11](#_Toc1144)

[6.1 问题分析 11](#_Toc7345)

[6.2 构造模型和求解示例 13](#_Toc18175)

[七、问题三的解 16](#_Toc10264)

[7.1 问题分析 16](#_Toc31218)

[7.2 模型构建与求解 17](#_Toc9571)

[八、问题四的解 22](#_Toc32450)

[8.1 问题分析 22](#_Toc25476)

[8.2 模型构建与求解 22](#_Toc5841)

[8.2.1回归分析算法 22](#_Toc31651)

[8.2.2主因分析算法 26](#_Toc9903)

[8.2.3构建对竞拍价预测的BP神经网络 28](#_Toc22513)

[8.2.4差分进化算法 30](#_Toc12327)

[九、问题五的解 31](#_Toc6661)

[参考文献 35](#_Toc8605)

[附录 35](#_Toc9519)

# 一、问题的重述

21世纪以来，科技发展迅速，尤其是媒体行业更是充斥着人们的生活。其中，电视台是发展较早的传统意义上的媒体，到现在已经基本覆盖了全国范围，而电视广告是地方电视台商业营运的主要业务之一。 因此电视台广告的推送需要好的方案：对产品销售或制造商来说，信息技术的发展要求提出更好的广告营销方案；对电视频道来说，不同的用户具有不同的特征，对不同的广告有不同的接受能力，进而影响广告商的收益。

因此我们设计和建立合适的模型进行预估：假设广告视频由产品销售或制造商（下统称买方）负责制作并向电视台（下称卖方）竞买播放时段。假设广告播放时段分周期组织竞卖，当前周期组织完成下一个周期的竞价交易。主要任务是：

1、通过选择/提取视频广告和电视频道用户的分类特征建立二者分类匹配推送静态模型；

2、设计卖方分时段竞卖合理底价估算模型；

3、基于已知电视频道用户收视历史情况和在播视频广告的产品销售情况，建立视频广告与频道用户的分类匹配推送更新模型；

4、基于更新的分户推送方式建立竞价交易模型，在极大化卖方收益的同时提升收视率和买方产品销售量；

5、设计并实现建立的模型的求解算法，通过收集或编撰数据给出算例。

# 二、问题分析

本文要求设计和建立合适的模型：假设广告视频由买方负责制作并向卖方竞买播放时段，广告播放时段分周期组织竞卖，当前周期组织完成下一个周期的竞价交易进行预估。所以得先考虑买方和卖方的成本和利益分配，以及其影响因素。

问题一中，要求建立二者分类匹配静态模型，可以找出各类用户与不同广告的匹配度。因此首先，我们将电视频道用户按年龄特征分为19-25岁，26-30岁，31-45岁，46岁-65岁以及女年轻人26-45岁。利用层次分析法建立层次结构模型，根据每一群体特征建立其与多种商品广告类型，建立二者分类匹配推送静态模型；

问题二，要求我们设计卖方分时段竞卖合理底价估算模型，在处理时分析得到现有主流媒体的收视率和底价之间的关系，加以考虑到卖方的主要分类和影响因素的不同：从观看电视节目的时段活跃度、电视台自身活跃指数、电视节目及其收视人群分布，以及节目收视人群经济实力为出发点给予每部分不同的权重，最后综合给出合适的底价；

问题三，在建立匹配机制时，由于人群的特征和广告的类型非常多，如果尝试建立一个方程的话，该方程就会非常复杂，所以基于在大数据的背景下，通过BP神经网络算法来对问题的求解进行简化。由于该算法在进行训练时，将人群的特征和广告的类型进行了特征提取，所以能提高了算法的匹配精确度。

问题四，在建立竞价交易模型时，我们需要保证竞价的正常进行，并且需要考虑多个影响因素，例如广告商的个数及他们各自的预算，电视台的底价以及收视率等许多方面，因此我们首先研究竞价过程能否开始，再对竞价的结果即最终成交价进行研究，并且注意到电视台的成本问题，例如宣传费用等，最终成交价减去成本即为电视台的收益问题。

问题五，我们需要收集合适的数据或编撰具有合理性、广泛性的数据，利用上述的数学模型进行求解实际问题，并且得到准确度较高的结果。

# 三、模型假设

1.假设广告视频由产品销售或制造商负责制作并向电视台竞买播放时段。

2.假设广告播放时段分周期组织竞卖，当前周期组织完成下一个周期的竞价交易。

3.假设每个时间段均有不同的广告商参与竞价。

# 四、符号约定

表4-1符号说明表

|  |  |  |  |
| --- | --- | --- | --- |
| 符号 | 含义 | 符号 | 含义 |
|  | 实用性 |  | 电视节目的时段活跃度 |
|  | 兴趣性 |  | 电视台自身活跃指数 |
|  | 经济性 |  | 电视节目及其收视人群分布 |
|  | 18-25岁 |  | 节目收视人群经济实力 |
|  | 26-30岁 |  | 底价和收视率系数 |
|  | 31-45岁 |  | 年龄 |
|  | 46-65岁 |  | 性别逻辑变量 |
|  | 女年轻人（26-45岁） |  | 广告观看逻辑变量 |
|  | 电子产品类 |  | 电视台的底价 |
|  | 金融类 |  | 广告商的预算 |
|  | 家电类 |  | 预期销售额 |
|  | 医疗养生类 |  | 电视台的收视率 |
|  | 服装类 |  | 广告商的个数 |
|  | 洗护用品类 |  | 电视台的宣传费用 |
|  | 汽车类 |  | 最终竞拍价 |
|  | 美妆类 |  | 电视台的收益 |
|  | 食品类 |  | 是否竞拍 |
|  | 母婴类 |  | 广告与用户的匹配度 |
|  | 特征向量或权向量 |  | 矩阵元素 |
|  | 一致性指标 |  | 平均随机一致性指标 |
|  | 随机一致性比率 |  | 第n类广告的销售情况 |
|  | 归一化后的销售情况 |  |  |

# 五、问题一的解

## 5.1 问题分析

分析题目要求，建立二者分类匹配静态模型，可以找出各类用户与不同广告的匹配度。因此首先，我们将电视频道用户按年龄特征分为19-25岁，26-30岁，31-45岁，46岁-65岁以及女年轻人26-45岁。我们的分类依据主要基于以下几个方面：

（1）如果按年龄将整个社会划分为学龄前儿童、学龄儿童、青少年、18至35岁的年轻人、35至65岁的中年人以及65岁的老年人等六类人群。但明显存在的问题是对于青少年以及儿童来说几乎没有经济能力，而65岁以上的老年人消费能力较低，因此本次分类仅考虑了19岁-65岁年龄段的群体，且在年龄分段上更加细致。

（2）具体的分类，我们试着从年龄属性和社会属性进行划分。消费群体可以分为五种：19-25岁、26-30岁、31-45岁、46-65岁，同时，因为女性年轻人（26-45岁）在年轻人消费中占据独特的位置，我们单独列出来分析。

19-25岁的群体特征：追求新颖，新潮流，往往因为款式、颜色、形状或价格等因素发生冲动性购买。但大多数仍没有经济独立，持续购买力低，因此相对来说在兴趣方面消费能力最强。

26-30岁的群体特征：在消费需求构成上，这个年龄阶段的人群大多结婚成家，有自己的子女，因此在需求构成和顺序上，家庭的需求数量最大，小孩的需求最重要，其次才是穿着和食品。

31-45岁的群体特征：因为消费心理的成熟，他们在购买产品的过程中更注重实际，以方便实用为主，要求提供方便、良好的环境条件和服务。

46-65岁的群体特征：因为生理机能的演变，这年龄阶段的消费者开始对健康更加关注，对健康食品和用品的需求量也大大增加，需求结构开始发生变化，对穿着以及其他奢侈品方面的支出大大减少，而对健身娱乐、特殊兴趣嗜好、旅游观光等方面的消费明显增加。

女性年轻人（26-45岁）：就目前而言，26-45岁的女性年轻人绝大多数处于就业状态，平时，她们既要负担工作，又要负担大部分家务劳动，因此，他们对日常生活的便利性和实用性要求较强；同时，女性年轻人的新型消费产品层出不穷，因此整体的消费能力相比较其他群体要大大不同。

其次，根据调查结果如下图所示，我们将广告类型分为电子产品类、金融类、家电类、医疗养生类、服装类、洗护用品类、汽车类、美妆类、食品类、母婴类。

图5-1 广告比例调查图

## 5.2 层次分析法构造模型

层次分析法（AHP）在对复杂的决策问题的本质、影响因素及其内在关系等进行深入分析的基础上，利用较少的定量信息使决策的思维过程数学化，是对难以完全定量的复杂系统作出决策的模型和方法。

**1.模型分析**

由问题分析可知，我们需要得出各类人群对各种广告的购买意愿即匹配程序，此处我们针对每一种广告进行独立分析，将各个年龄段群体作为层次分析法的基础层，将此时分析的该种广告商品的实用性、经济性、兴趣性作为中间层，将各年龄段群体对广告商品的购买意愿作为目标层，即最终得到的权重最大的年龄段群体，即为与该广告商品匹配度最高的用户。

**2.模型建立**

运用层次分析法构造系统模型时，大体可以分为以下四个步骤：

（1）建立层次结构模型，如下图：

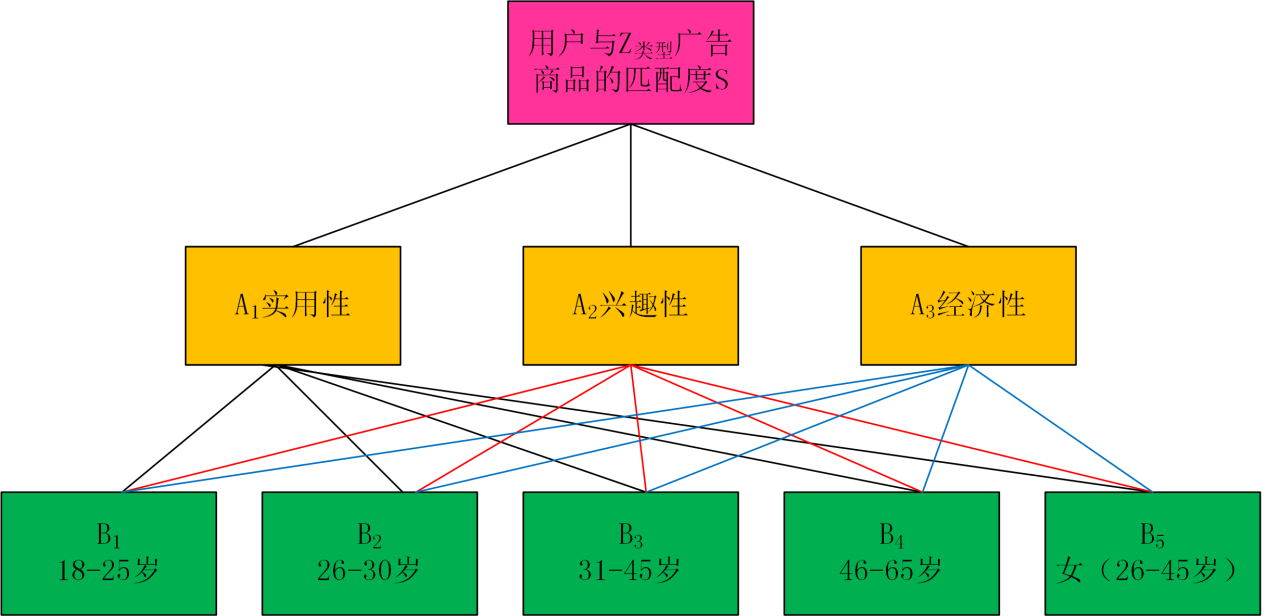


图5-2 层次模型结构图

其中绿色框图代表基础层，黄色框图为中间层，粉色框图为目标层。同时，我们设定用对应表示实用性、兴趣性、经济性，用对应表示18-25岁、26-30岁、31-45岁、46-65岁及女年轻人（26-45岁），用表示匹配度。

此外我们用，下标按顺序1-10分别表示电子产品类、金融类、家电类、医疗养生类、服装类、洗护用品类、汽车类、美妆类、食品类、母婴类广告。

（2）构造判断（此处为成对比较）矩阵：

在确定各个层次各因素之间的权重时，我们不能仅仅定性分析，采用Santy等人提出的一致矩阵法，即：不把所有因素放到一起比较，而是两两相互比较；此时采用相对尺度，以尽可能减少性质不同的诸因素相互比较的困难，以提高准确度。成对比较矩阵是表示本层所有因素针对上一层某一个因素的相对重要性的比较。其中成对比较矩阵的元素我们采用Santy的1-9标度方法给出，如下表所示：

表5-1 矩阵的元素Santy的1-9标度方法

|  |  |
| --- | --- |
| 标度 | 含义 |
| 1 | 表示两个元素相比，具有同样重要性 |
| 3 | 表示两个元素相比，前者比后者稍重要 |
| 5 | 表示两个元素相比，前者比后者明显重要 |
| 7 | 表示两个元素相比，前者比后者强烈重要 |
| 9 | 表示两个元素相比，前者比后者极端重要 |
| 2,4,6,8 | 表示上述相邻判断的中间值 |
| 倒数 | 若元素与的重要性之比为，那么元素与元素重要性之比为 |

我们首先构造基础层针对中间层因素的成对比较矩阵，如下：







以下我们以电子广告为例进行计算，由上面的假设可知电子广告为，根据客观情况下电子广告对用户兴趣性的要求最高，其次实用性，最后经济性，可以得出对的权重矩阵为：

（3）一致性检验；

在构造成对比较矩阵时，我们会发现出现成对比较的不一致情况，例如：

 ，由元素关系我们可以发现，，按正常一致比较情况下，应得，但矩阵中表示的是，出现了不一致情况。这种情况是很容易出现的，这是由于客观事物的复杂性和人的认识的多样性所决定的，因此我们允许不一致情况，但如果出现“甲比乙极端重要，乙比丙极端重要，而丙又比甲极端重要”这样的情况，是明显违反客观规律的，因此我们确定不一致的允许范围，本次设定为0.1。

（4）下面给出层次分析法计算的主要步骤：

①计算判断矩阵的每一行元素乘积，即

②计算的次方根，即

③若标准化为则为所求特征向量。

④计算最大特征值，式中表示向量的第个分量。

⑤计算判断矩阵的一致性指标，即，为矩阵的阶数。

⑥给出判断矩阵的同阶平均随机一致性指标，本次由已知资料搜索给出。

⑦利用上述判断矩阵的一致性指标与同阶平均随机一致性指标的比值得到随机一致性比率，记为，当其小于给出的0.1时，我们认为判断矩阵具有满意的一致性，否则就需要调整判断矩阵，使其具有满意的一致性。

当我们通过上述步骤得到基础层对于中间层的权向量，，以及中间层对于目标层的权向量，因此我们可以根据权重计算公式得到分别对的权值即我们所求的匹配度。例如，基础层对目标层的公式如下：



## 5.3 模型求解

通过MATLAB编写层次分析法程序，进而计算得到对的权重值为；对的权重值为；对的权重值为。

同时也通过MATLAB计算得到对的权重值为。

根据以上权重值及匹配度计算公式，我们可以得出电子广告与不同年龄段的用户之间的匹配度，同理经过建立与其他不同广告之间的成对比较矩阵，计算可以得到结果表格如下：

表5-2 不同广告与用户匹配表

| 广告类型 | 18-25岁 | 26-30岁  (男) | 31-45岁  (男) | 女年轻人  (26-45岁) | 46-65岁 |
| --- | --- | --- | --- | --- | --- |
| 电子产品 | 0.30 | 0.15 | 0.14 | 0.28 | 0.14 |
| 金融 | 0.07 | 0.14 | 0.25 | 0.28 | 0.25 |
| 家电 | 0.08 | 0.14 | 0.25 | 0.28 | 0.25 |
| 医疗养生 | 0.07 | 0.14 | 0.26 | 0.28 | 0.25 |
| 服装 | 0.30 | 0.15 | 0.14 | 0.28 | 0.14 |
| 洗护用品 | 0.10 | 0.14 | 0.24 | 0.28 | 0.24 |
| 汽车 | 0.07 | 0.14 | 0.25 | 0.27 | 0.26 |
| 美妆 | 0.16 | 0.14 | 0.21 | 0.28 | 0.21 |
| 美食 | 0.34 | 0.15 | 0.12 | 0.28 | 0.12 |
| 母婴 | 0.08 | 0.14 | 0.25 | 0.28 | 0.25 |

进行归一化处理后，得到以下表格

表5-3 不同广告与用户匹配归一化表

| 广告类型 | 18-25岁 | 26-30岁  (男) | 31-45岁  (男) | 女年轻人  (26-45岁) | 46-65岁 |
| --- | --- | --- | --- | --- | --- |
| 电子产品 | 1 | 0.5 | 0.47 | 0.93 | 0.47 |
| 金融 | 0.25 | 0.5 | 0.89 | 1 | 0.89 |
| 家电 | 0.29 | 0.5 | 0.89 | 1 | 0.89 |
| 医疗养生 | 0.25 | 0.5 | 0.93 | 1 | 0.89 |
| 服装 | 1 | 0.5 | 0.47 | 0.93 | 0.47 |
| 洗护用品 | 0.36 | 0.5 | 0.86 | 1 | 0.86 |
| 汽车 | 0.26 | 0.52 | 0.93 | 1 | 0.96 |
| 美妆 | 0.57 | 0.5 | 0.75 | 1 | 0.75 |
| 美食 | 1 | 0.44 | 0.35 | 0.82 | 0.35 |
| 母婴 | 0.29 | 0.5 | 0.89 | 1 | 0.89 |

归一化处理之后，我们可以更加直观的得到不同年龄段的用户和不同类型广告的匹配度，数值为1代表某种广告最适合的某年龄段用户，随着数值减小代表适合度逐渐减小。

根据日常生活规律，我们假设某时间段各年龄段的人群所占的比例分别是：18-25岁占15%，26-30岁（男）占25%，31-45岁（男）占20%，26-45岁（女）占30%，46-65岁占10%，总人数为一万人，并且当前时间段的各类型广告平均分布。根据上述得到的不同广告与用户匹配归一化表，我们可以得到这一时间段各类广告商品的销售量，如下表：

表5-4 不同广告与用户匹配销售量表

| 广告类型 | 18-25岁 | 26-30岁  (男) | 31-45岁  (男) | 女年轻人  (26-45岁) | 46-65岁 | 广告商品销售量 |
| --- | --- | --- | --- | --- | --- | --- |
| 电子产品 | 1500 | 1250 | 940 | 2790 | 470 | 6950 |
| 金融 | 375 | 1250 | 1780 | 3000 | 890 | 7295 |
| 家电 | 435 | 1250 | 1780 | 3000 | 890 | 7355 |
| 医疗养生 | 375 | 1250 | 1860 | 3000 | 890 | 7375 |
| 服装 | 1500 | 1250 | 940 | 2790 | 470 | 6950 |
| 洗护用品 | 540 | 1250 | 1720 | 3000 | 860 | 7370 |
| 汽车 | 390 | 1300 | 1860 | 3000 | 960 | 7510 |
| 美妆 | 855 | 1250 | 1500 | 3000 | 750 | 7355 |
| 美食 | 1500 | 1100 | 700 | 2460 | 350 | 6110 |
| 母婴 | 435 | 1350 | 1780 | 3000 | 890 | 7455 |

可以看出，各类广告商品的销售量存在一定的差异，但基本满足市场规律。此外，对商品销售量影响最小的是18-25岁的人群和46-65岁的人群，除去人数比例较少的影响，也与18-25岁人群经济能力不足、46-65岁人群对商品需求少、二者购买力都比较弱有关。通过上述假设，进一步说明了不同用户与广告的匹配度是符合客观规律的。

# 六、问题二的解

## 6.1 问题分析

通过收集主要电视频道的广告刊例进行分析，发现广告刊例的涨幅与电视节目的实时收视率呈现出极大的相关性，收视率是指某一时段内收看某一节目的人数占电视观众总人数的百分比（收视率的计算方法：收视率=观看该电视节目人数/总人数\*100%）。而电视节目的收视率的影响因素与各个家庭的背景资料数据，包括性别、年龄、职业、文化程度、个人/家庭收入、个人/家庭资产、家庭规模、家庭角色关系、家庭收视设备占有情况和电视接收情况、基本生活态度和休闲选择、工作情况、居住地区、使用语言等。

电视收视率指标被区分为两大类，一类为基本指标，指各种计算口径的收视率，如按人计算的收视率、按户计算的收视率、按时段计算的收视率、按节目计算的收视率、按目标观众计算的收视率等。另一类为派生指标，包括反映市场结构的指标如收视份额（或称市场占有率）、观众构成等，反映节目收视量的指标如开机率、总收视点（grps）、平均收视时间、到达率和暴露频次等，反映节目收视偏好的指标如节目吸引力、观众重叠率、观众忠诚度等。这些指标有的更受电视台重视，因为其对指导电视节目编排更有意义；有的更受广告公司和广告主重视，因为其对测算广告投放成本、优化广告投放策略更有意义。例如ggtam中分别对总收视点、开机率、到达率、暴露频次、收视份额等主要指标定义如下： 总收视点（grps：gross rating points）：某项广告投放活动中各次插播点收视率的总和； 开机率（huts/puts：household/people using television）：某特定时间段内按户或按人的电视平均开机（不论看何频道）百分比； 到达率（reach）：作为观众收看某一系列节目或某一广告排期插播点至少一次的累积百分比； 暴露频次（frequency）：一个广告排期或一系列节目中观众的平均收看次数； 收视份额（audience share）：某一规定时段内，某特定频道或节目的观众收视量占观众总收视量的百分比。

分析得到这些数据额需要对背景人群进行抽样调查，调查数据包括两部分：

一部分是被调查者（被调查个人及其所属户）的背景资料数据，包括性别、年龄、职业、文化程度、个人/家庭收入、个人/家庭资产、家庭规模、家庭角色关系、家庭收视设备占有情况和电视接收情况、基本生活态度和休闲选择、工作情况、居住地区、使用语言等。这部分数据主要通过样户入网调查得到，成为进行收视率计算和分析的背景数据。

另一部分是通过日记法或人员测量仪法记录到的被调查者个人的电视收视数据，具体即有电视的家庭户（样本户）中每位成员（年龄4周岁及以上的被调查者）在每天连续的24小时内每15分钟（人员测量仪为每分钟）收看或不收看电视以及收看什么频道的记录信息。

另外还有一部分数据不属于收视率调查数据，也就是说不是从样户调查得来的数据，但对收视率计算与分析也至关重要，这就是电视频道属性数据和电视节目播出数据。这部分数据一般由电视台提供或通过电视监播得到。

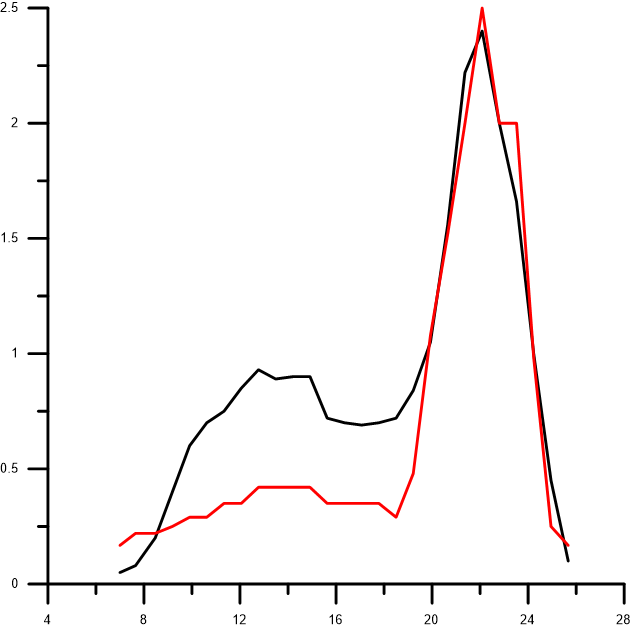
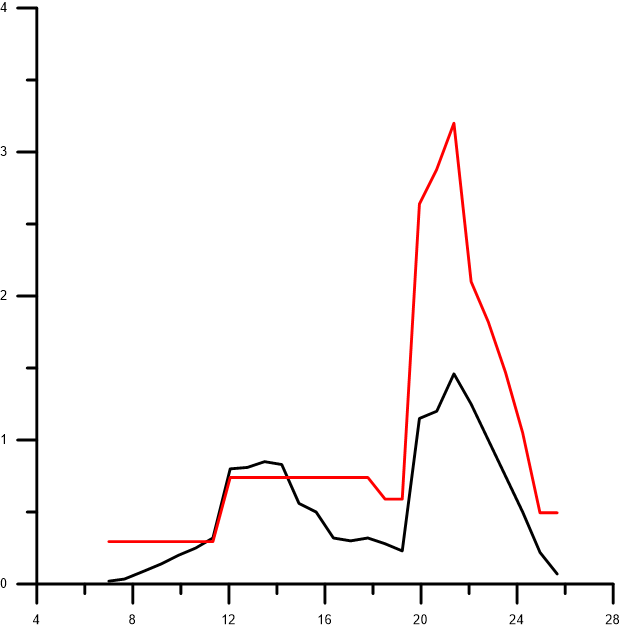
有了以上三部分数据，就使得收视率指标的计算成为可能。只根据电视收视数据可以计算总收视率及分时段分频道的收视率；如果结合背景资料数据，则可计算分目标观众收视率；如果结合电视节目播出数据，则可计算节目收视率。这样组合运用这些数据，就可以计算出一系列的收视率指标。

## 6.2 构造模型和求解示例

为更好地满足电视媒体、广告行业对收视调查服务的需求，央视－[索福瑞](https://www.baidu.com/s?wd=%E7%B4%A2%E7%A6%8F%E7%91%9E&tn=SE_PcZhidaonwhc_ngpagmjz&rsv_dl=gh_pc_zhidao" \t "https://zhidao.baidu.com/question/_blank)媒介研究有限公司(CSM)，对全国近700个主要电视频道的收视情况进行全天候不间断的监测。

以主要中国主要活跃媒体为例：

湖南卫视、江苏卫视、浙江卫视、北京卫视、东方卫视：

  图6-1 湖南卫视收视率-底价实际图 图6-2 江苏卫视收视率-底价实际图

时间/h

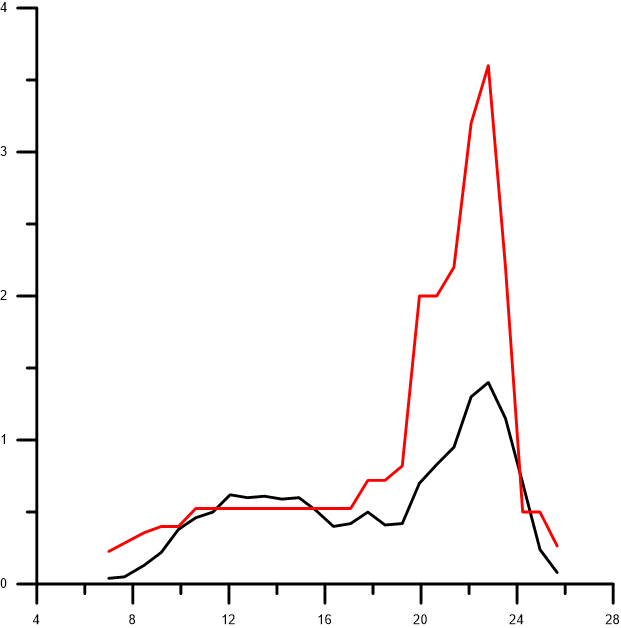
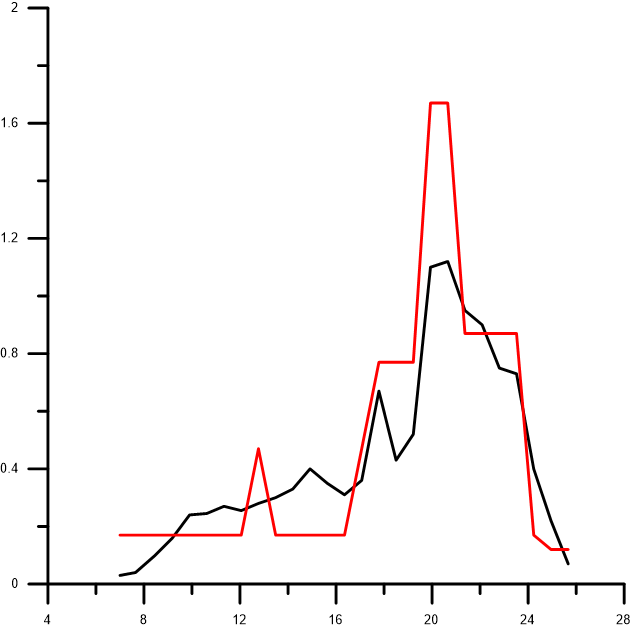
收视率

广告位/十万

收视率

广告价位/十万

时间/h

收视率

广告价位/十万

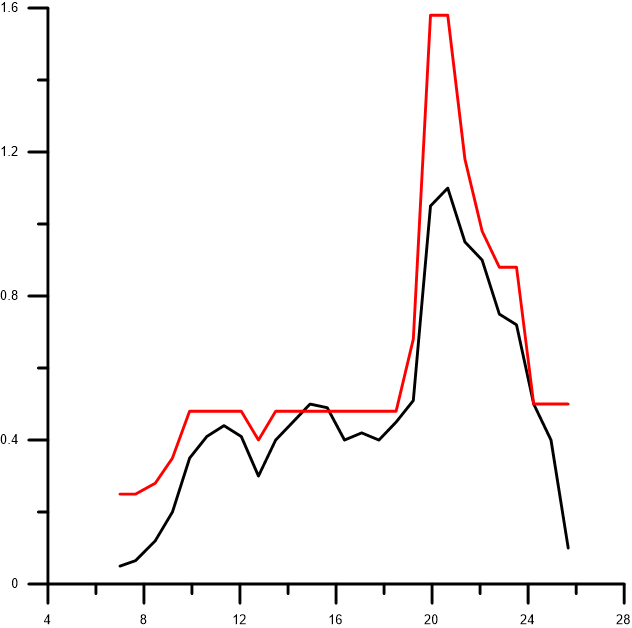
时间/h

收视率

广告价位/十万

时间/h

图6-3 浙江卫视收视率-底价实际图 图6-4 北京卫视收视率-底价实际图



收视率

广告价位/十万

时间/h

图6-5 东方卫视收视率-底价实际图

通过对比可知，收视率和广告价位确实基本成比例关系，所以可以得出，根据电视台的收视率指标可以来定义广告价位的标准，基本呈现线性关系；所以我们认为，收视率和广告价位可以用线性拟合，因此根据收视率来建立模型。

收视率的计算我们以性别、年龄、收入、电视收视率情况为主要考虑因素，因为前三个因素各个类别可以代表，某一群体在观看电视节目时的时间段、兴趣爱好、购买能力、通过与收视率的结合便可以计算出：这一群体对某一货物或商品的潜在购买力度，对不同的商品广告，不同人群的接受度不同，以及不同的购买心理反馈。

对此我们赋予不同影响因素，不同的权重，

①底价=K\*收视率

②收视率= P1\*时段活跃指数+P2\*电视台活跃指数+P3\*节目内容（收视人群基数）+P4\*节目收视人群经济实力（P1、P2、P3、P4为不同的权重，在模拟时段：P1=0.5；P2=0.25；P3=0.15；P4=0.1）

通过对比分析得到，底价和收视率之间的线性比为100000

③底价=100000\*（0.5\*时间活跃指数+0.25\*电视台活跃指数+0.15\*节目内容（收视人群基数）+0.1\*节目收视人群经济实力

通过这样的模型计算给出某一产品适合的播放时段和其对应的底价分别是多少。

在此之前进行背景调查和分析：以1000人为样本，进行抽样分析，可以得到：

表6-1背景调查表

|  |  |  |  |
| --- | --- | --- | --- |
| 年龄 | 观看时间时段 | 主要关注节目类型 | 主要观看哪些电视台 |
| 18~25 | 中午12点至一点  下午7点至晚上12点  （有一部分到晚上1点左右） | 综艺、电视剧（IP）、电影、体育节目 | 湖南、江苏、东方等 |
| 26~30 | 下午8点至晚上12点 | 电视剧（热门）和综艺 | 湖南等热门为主 |
| 31~45 | 晚上9点至11点 | 电视剧（历史和家庭） | 不在意电视台的区别 |
| 46~60 | 早上7点至8点  晚上6:30至8:30 | 新闻、天气预报 | 中央电视台为主 |

在此过程中存在问题：

男士和女士感兴趣的节目并不相同，未考虑男女比例，所以对男士的较为感兴趣的阶段，比如体育节目播放时段和频道到的底价做出了不恰当的预估，所以加入观看电视节目男女比例：

图6-6 观众分布图

综上所述，列出各项对应的收视率（底价）的参考模型；

输入时段、电视台；

②、数据库中进行比较迭代,找到该电视台与该时段相距最近的，最多播出的内容和观众人群指数；

③、将②的四个指数代入底价计算公式

④、得出该时段的预计估价。

表6-2底价模型算例表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 输入 | | 参数 | | | | 输出 |
| 时间 | 电视台 | 时段活  跃指数 | 电视台活跃指数 | 节目内容（收视人群基数） | 节目收视人群经济实力 | 底价 |
| 16:00 | 东方 | 0.25 | 0.5 | 0.25 | 3 | 58750 |
| 18:00 | 北京 | 0.5 | 0.75 | 0.5 | 3 | 71250 |
| 20:00 | 浙江 | 1.5 | 1.5 | 1.5 | 7 | 205000 |
| 22:00 | 湖南 | 2 | 2 | 1.5 | 7 | 246250 |
| 24:00 | 江苏 | 0.75 | 1.5 | 0.75 | 6 | 146250 |

表6-3底价模型与实际底价误差对比表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 电视台 | 东方 | 北京 | 浙江 | 湖南 | 江苏 |
| 时间段 | 16:00 | 18:00 | 20:00 | 22:00 | 24:00 |
| 估算底价 | 58750 | 81250 | 205000 | 246250 | 146250 |
| 实际底价 | 61000 | 77000 | 200000 | 250000 | 147000 |
| 误差百分比 | -3.69% | 5.52% | 2.50% | -1.50% | -0.51% |

# 七、问题三的解

## 7.1 问题分析

对于问题二，主要是分成2个步骤来进行解答的：

首先获取各个用户的收视历史情况和各类商业广告的销售情况，根据第一问的分类方法，提取各个用户的特征（即年龄和性别），再提取商业广告的种类（即第一种方法划分的10类广告类型），这样就可以得到一个的关于特征性的矩阵，其中n代表个体数，此外再根据各个个体对这10类商业广告的购买情况，可得到一个销售量为的向量；

利用BP神经网络进行各类特征对于销售量的影响情况的分析；其中关于特征性的矩阵为输入，的销售量向量为目标数据，通过训练得到一个关于各个特征属性和销售量关系的神经网络；

计算各类人群属性特征对各类广告的观看后的销售关系。



图7-1 人群属性和各类广告观后反馈图

## 7.2 模型构建与求解

一个商业广告的产品销量不仅与商品自身有光，一般还与受众的兴趣、经济和收视情况有关，所以在分析销售量的影响时，应该对人群和广告进行特征的提取。

1.根据第一问的分类方法，人群类型的划分主要分成年龄，此外由于女年轻人的特殊性，所以单独列了出来，为了方便列出矩阵，所以将性别也当作人群分类的特征；而广告类型同第一问一样，不同的广告类型即作为广告的特征；所以从各个个体的信息中，我们可以提取得到以下的类型：

[年龄 性别 电子产品类广告 ..... 母婴类广告]

上面的类型即为关于特征的输入信息矩阵，而目标结果向量则为个体的各类广告的销售情况，所以可得到矩阵形式：





是第n类广告的销售情况

age是每个个体的年龄大小，由于本文在考虑人群时只考虑了岁的，所以在年龄的输入上进行归一化处理：



表7-1逻辑变量声明

| 逻辑变量 | 男 | 女 |
| --- | --- | --- |
|  | 0 | 1 |
| 逻辑变量 | 未观看过 | 观看过 |
|  | 0 | 1 |

2、 BP神经网络计算

BP神经网络的学习过程主要由四部分组成：输入模式顺传播、输出误差逆传播、循环记忆训练、学习结果判别。在正向传播过程中，输入信息从输入层经隐含层单元逐层处理，并传向输出层，每一层神经元的状态只影响下一层神经元的状态。如果在输出层不能得到所期待的输出，则转入反向传播，将误差信号沿原来的连接通路返回，通过修改各层神经元的权值，使得误差信号减小，然后再转入正向传播过程。反复迭代，直到误差小于给定的值为止。



图7-2 神经网络流程图

3、各类特征对销售量的影响

首先控制输入变量得到各个年龄对不同广告类型的销售情况，并进行了归一化：



Q为归一化后的销售情况。下面分别是按图表7-2顺序表示的广告类型对不同年龄的匹配度图。

图片包含 地图

描述已自动生成

图片包含 地图, 文字

描述已自动生成

图片包含 地图

描述已自动生成图片包含 地图, 文字

描述已自动生成

图片包含 地图

描述已自动生成图片包含 地图

描述已自动生成



图7-3 归一化的销售情况图

将上面各个年龄段之内的销售情况求平均即可得到我们下要的各个年龄段的销售情况：销售情况如下表：

表7-2 销售情况分布表

| **广告类型** | **18-25岁** | **26-30岁**  **(男)** | **31-45岁**  **(男)** | **女年轻人**  **(26-45岁)** | **46-65岁** |
| --- | --- | --- | --- | --- | --- |
| 电子产品 | 0.93 | 1.0 | 0.78 | 0.80 | 0.14 |
| 金融 | 0.06 | 0.24 | 0.61 | 0.37 | 0.94 |
| 家电 | 0.20 | 0.57 | 0.87 | 0.72 | 0.11 |
| 医疗养生 | 0.10 | 0.28 | 0.94 | 0.88 | 0.86 |
| 服装 | 0.79 | 0.70 | 0.33 | 0.88 | 0.04 |
| 洗护用品 | 0.51 | 0.85 | 0.87 | 0.84 | 0.06 |
| 汽车 | 0.07 | 0.32 | 0.84 | 0.36 | 0.4 |
| 美妆 | 0.45 | 0.26 | 0.05 | 0.92 | 0.34 |
| 美食 | 0.88 | 0.64 | 0.35 | 0.59 | 0.04 |
| 母婴 | 0.12 | 0.44 | 0.66 | 0.82 | 0.14 |

最后，观测表格数据可以发现，不同的广告类型对不同人群的影响是非常大的，从电子产品而言，岁之间的购买意向是最大的，根据第一问的分析资料也可得到，该年龄段正好处于对新型事物感兴趣和喜爱追求新潮流的阶段，并且还具有一定的经济基础，所以相对于岁之间的稍微大一点；同时，我们可以发现，当为母婴广告时，岁的销售情况变得很小，对于低年龄段而言，由于晚婚晚育的影响，所以对母婴产品的需要特别少，而高年龄段则由于过了生育年龄，所以对母婴产品的需求也会快速下降。所以对于不同类型广告的不同年龄的划分对销售量的影响是可行的。

# 八、问题四的解

## 8.1 问题分析

分析问题四的要求，建立竞价交易模型，极大化卖方的收益的同时提升收视率和买方产品销售量，这样的交易模式在日常生活中处处可见，经过前期工作的调查及研究，我们针对竞价交易模型从以下两个方面进行研究：

①某播放时段是否竞拍，主要受电视台给出的底价、电视台的收视率、买方的预算以及买方预期的销售额四个方面影响；

②另外一方面是对竞拍价格的影响主要包括电视台给出的底价、电视台的收视率、买方的预算、买方预期的销售额、买方的个数以及电视台的宣传费用等因素。

以下我们设电视台给出的底价为，买方预期的销售额为，买方的预算为，电视台的收视率为，买方的个数为，电视台的宣传费用为，最终竞拍价为，电视台的收益为。

## 8.2 模型构建与求解

通过问题分析，我们以下就从上述两个方面建立数学模型即并分为（1）（2）两部分进行模型建立。

### 8.2.1回归分析算法

回归分析是一种统计学上分析数据的方法，可以有效帮助我们了解两个或多个变量间是否相关、相关方向与强度，并建立数学模型以便观察特定变量来预测研究者感兴趣的变量。

虽然自变量和因变量之间没有严格的、确定的函数关系，但可以设法找出最能代表它们之间关系的数学表达形式。例如：一元线性回归方程为：



本次研究选用了多元回归分析，多元回归分析实质上就是在自变量很多时，其中有的因素可能对因变量的影响不是很大，而且变量之间可能不完全相互独立的，可能有种种互相作用的关系。在这种情况下，采用逐步回归分析将会根据预先设定的统计量的概率值进行筛选，此时最先进入回归方程的自变量应该是跟因变量关系最为密切、贡献最大的，这样建立的多元回归模型预测结果会更好。那么我们最终建立的多元线性回归方程应该为：



其中代表随机误差，其中随机误差分为：可解释的误差和不可解释的误差，随机误差必须满足以下四个条件，多元线性方程才有意义：

①服从正态分布，即指随机误差必须是服从正态分布的随机变量；

②无偏性假设，即指期望值为0；

③同共方差性假设，即指所有的随机误差变量方差都相等；

④独立性假设，即指所有的随机误差变量都相互独立，可以用协方差解释。

接下来进行模型求解，利用录入好的Excel数据文件和SPSS分析软件，我们将数据文件导入，并根据下表确定自变量和因变量

表8.1 自变量和因变量

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 因变量 | 自变量 | | | |
| 是否竞拍P | 电视台给出的底价为 | 买方预期的销售额为 | 买方的预算为 | 电视台的收视率为 |

注：P=1即为存在竞拍，P=0即为不竞拍。

SPSS多元线性回归分析的结果如下所示并进行结果分析：

|  |  |  |  |
| --- | --- | --- | --- |
| 表8.2 输入／移去的变量a | | | |
| 模型 | 输入的变量 | 移去的变量 | 方法 |
| 1 | 广告商预算 | . | 步进（准则: F-to-enter 的概率 <= .050，F-to-remove 的概率 >= .100）。 |
| 2 | 底价 | . | 步进（准则: F-to-enter 的概率 <= .050，F-to-remove 的概率 >= .100）。 |
| 3 | 预期销售额 | . | 步进（准则: F-to-enter 的概率 <= .050，F-to-remove 的概率 >= .100）。 |
| 4 | 收视率 | . | 步进（准则: F-to-enter 的概率 <= .050，F-to-remove 的概率 >= .100）。 |

由于选项中选择的是“逐步法”，逐步法是向前和向后的结合体，从结果可以看出，最先进入“线性回归模型的”是“广告商预算”，其次是“底价”，然后是“预期销售额”，最后是“收视率”。此外，设置的进入和删除的概率值为0.05和0.10，表示当小于等于0.05时，进入 “线性回归模型”，当大于等于0.1时，从“线性模型”中剔除。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 表8.3 模型汇总 | | | | |
| 模型 | R | R方 | 调整R方 | 标准估计的误差 |
| 1 | .512a | .263 | .263 | .429 |
| 2 | .718b | .516 | .516 | .348 |
| 3 | .718c | .516 | .516 | .348 |
| 4 | .718d | .516 | .516 | .348 |
| a. 预测变量: (常量)，广告商预算。 | | | | |
| b. 预测变量: (常量)，广告商预算，底价。 | | | | |
| c. 预测变量: (常量)，广告商预算，底价，预期销售额。 | | | | |
| d. 预测变量: (常量)，广告商预算，底价，预期销售额，收视率。 | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 表8.4 Anovaa | | | | | | |
| 模型 | | 平方和 | df | 均方 | F | Sig. |
| 1 | 回归 | 6563.066 | 1 | 6563.066 | 35603.836 | .000b |
| 残差 | 18433.225 | 99998 | .184 |  |  |
| 总计 | 24996.291 | 99999 |  |  |  |
| 2 | 回归 | 12889.537 | 2 | 6444.769 | 53231.238 | .000c |
| 残差 | 12106.754 | 99997 | .121 |  |  |
| 总计 | 24996.291 | 99999 |  |  |  |
| 3 | 回归 | 12893.329 | 3 | 4297.776 | 35508.700 | .000d |
| 残差 | 12102.962 | 99996 | .121 |  |  |
| 总计 | 24996.291 | 99999 |  |  |  |
| 4 | 回归 | 12894.462 | 4 | 3223.615 | 26636.090 | .000e |
| 残差 | 12101.830 | 99995 | .121 |  |  |
| 总计 | 24996.291 | 99999 |  |  |  |
| a. 因变量：是否竞拍。 | | | | | | |
| b. 预测变量：(常量)，广告商预算。 | | | | | | |
| c. 预测变量：(常量)，广告商预算，底价。 | | | | | | |
| d. 预测变量：(常量)，广告商预算，底价，预期销售额。 | | | | | | |
| e. 预测变量：(常量)，广告商预算，底价，预期销售额，收视率。 | | | | | | |

根据上述“模型汇总”表格可以看出，有四个模型，其中从的拟合程度来看，可以看出，模型2、3、4的拟合程序明显比模型1要好。

根据“Anova”表格可以看出模型2、3、4的回归平方和与残差平方和比较接近，由于总平方和=回归平方和+残差平方和，而残差平方和即为随机误差（不可解释的误差），所以此线性回归模型只解释了总平方和的一半。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 表8.5 系数a | | | | | | |
| 模型 | | 非标准化系数 | | 标准系数 | t | Sig. |
| B | 标准误差 | 试用版 |
| 1 | (常量) | .228 | .002 |  | 116.691 | .000 |
| 广告商预算 | .014 | .000 | .512 | 188.690 | .000 |
| 2 | (常量) | .499 | .002 |  | 252.202 | .000 |
| 广告商预算 | .027 | .000 | .994 | 326.259 | .000 |
| 底价 | -.027 | .000 | -.696 | -228.592 | .000 |
| 3 | (常量) | .499 | .002 |  | 252.262 | .000 |
| 广告商预算 | .025 | .000 | .920 | 67.657 | .000 |
| 底价 | -.027 | .000 | -.696 | -228.654 | .000 |
| 预期销售额 | .002 | .000 | .075 | 5.597 | .000 |
| 4 | (常量) | .496 | .002 |  | 214.428 | .000 |
| 广告商预算 | .025 | .000 | .919 | 67.649 | .000 |
| 底价 | -.027 | .000 | -.706 | -164.513 | .000 |
| 预期销售额 | .002 | .000 | .075 | 5.608 | .000 |
| 收视率 | 007 | .002 | .011 | 3.059 | .002 |

从“系数a”表中可以看出广告商预算和底价对因变量是否竞拍影响最大，且广告商预算与是否竞拍为正相关关系，底价与为负相关关系，且预期销售额和收视率对是否竞拍影响较小，利用非标准系数确定的多元线性回归方程应该为，利用标准系数确定的多元线性回归方程应该为，可以看出各自变量对因变量的影响大小排序为。由于我们考虑的是实际生活中的竞价问题，所以我们采用非标准系数线性回归方程。

在得到回归方程的基础上，为保证本次研究的准确性，我们抽取样本各1000个，代入到回归方程中，设定以下关系式：

其中为阈值。

我们队的阈值进行多次取值试验，最终得到当时，满足上述回归方程的样本数的饼状图如下所示：

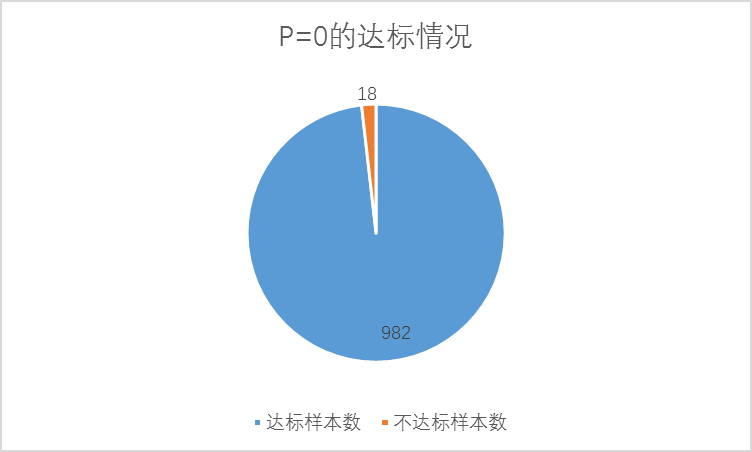


图8.1 样本达标率

### 8.2.2主因分析算法

因子分析法是从研究变量内部相关的依赖关系出发，把一些具有错综复杂关系的变量归结为少数几个综合因子的一种多变量统计分析方法。它的基本思想是将观测变量进行分类，将相关性较高，即联系比较紧密的分在同一类中，而不同类变量之间的相关性则较低，那么每一类变量实际上就代表了一个基本结构，即公共因子。对于所研究的问题就是试图用最少个数的不可测的所谓公共因子的线性函数与特殊因子之和来描述原来观测的每一分量。

在前面已经提到，在竞价模型中，我们首先讨论是否会发生竞拍的情况，然后后面的研究则是在确保竞拍的情况下，竞拍价是如何变化的。首先，根据市场情况分析和网上资料搜索分析，可以得到对竞拍价的影响因素比较多，如果直接直接进行分析，会显得过于复杂，所以，采用因子分析法进行降维处理，得到少量的主因变量。其中因变量和自变量为：

表8.6 自变量和因变量

| **因变量** | **自变量** | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| 竞拍价 | 竞拍底价 | 广告商预算 | 广告商预期销售额 | 收视率 | 广告商数量 | 竞拍宣传费 |

将相关自变量导入到SSPS软件中，进行主因分析，可得到：

表8.7 KMO 和 Bartlett 的检验

| **KMO 和 Bartlett 的检验** | | |
| --- | --- | --- |
| 取样足够度的 Kaiser-Meyer-Olkin 度量 | | .756 |
| Bartlett 的球形度检验 | 近似卡方 | 87786.807 |
| df | 15 |
| Sig. | .000 |

表8.8 KMO值检验表

|  |  |  |
| --- | --- | --- |
| 检测类别 | 值的范围 | 因子分析适合情况 |
| KMO值 | 大于0.9 | 非常适合 |
| 0.8~0.9 | 很适合 |
| 0.7~0.8 | 适合 |
| 0.6~0.7 | 不太适合 |
| 0.5~0.6 | 勉强适合 |
| 小于0.5 | 不适合 |
| BartlettP值 | 小于或者等于0.01 | 适合 |

从上表可以得到本文的因子分析的KMO=0.756，是属于合适的区间，并且Sig=.000，Sig值小于0.005，符合标准，数据呈球形分布，各个变量在一定程度上相互独立。综合来说，说明变量之间有共享因子的可能性，适合进行因素分析。

表8.9 各变量的相关矩阵

|  | | 商家个数 | 宣传费用 | 收视率 | 销售额 | 底价 | 最高报价 |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 相关 | 商家个数 | 1.000 | .960 | .005 | .009 | .005 | .009 |
| 宣传费用 | .960 | 1.000 | .006 | .006 | .004 | .008 |
| 收视率 | .005 | .006 | 1.000 | .762 | .807 | .769 |
| 预期销售额 | .009 | .006 | .762 | 1.000 | .945 | .965 |
| 底价 | .005 | .004 | .807 | .945 | 1.000 | .953 |
| 广告商预算 | .009 | .008 | .769 | .965 | .953 | 1.000 |

从上表可以分析得到，商家的个数和宣传费用是有很大的相关性的，但与其他的4个变量的相关性特别小，而收视率、预期销售额、底价以及广告商预算是有很大的相关性的，与其他2个变量关联很小，所以可以认为主要是由2个不关联的因子影响竞拍价。

表8.10 解释的总方差

| 成份 | 初始特征值 | | | 提取平方和载入 | | |
| --- | --- | --- | --- | --- | --- | --- |
| 合计 | 方差的% | 累积% | 合计 | 方差的% | 累积% |
| 1 | 3.608 | 60.125 | 60.125 | 3.608 | 60.125 | 60.125 |
| 2 | 1.960 | 32.669 | 92.794 | 1.960 | 32.669 | 92.794 |
| 3 | .306 | 5.097 | 97.891 |  |  |  |
| 4 | .053 | .889 | 98.779 |  |  |  |
| 5 | .040 | .662 | 99.442 |  |  |  |
| 6 | .034 | .558 | 100.000 |  |  |  |
| 提取方法：主成份分析。 | | | | | | |

从上表来看，当因子数为2时，累计方差达到了92.794%，后面的因子影响特别小，，所以利用因子分析法提取出2个因子是比较合理的。

表8.11 成份得分系数矩阵表

|  | 成份 | |
| --- | --- | --- |
| 1 | 2 |
| 商家个数 | .004 | .505 |
| 宣传费用 | .004 | .505 |
| 收视率 | .242 | -.004 |
| 预期销售额 | .269 | -.004 |
| 底价 | .271 | -.005 |
| 广告商预算 | .270 | -.003 |
| 提取方法 :主成份分析。 | | |

根据上表的成份得分系数矩阵表，可以得到2个因子的关系上为：





### 8.2.3构建对竞拍价预测的BP神经网络

从上面通过因子分析后，可以得到2个因子，再将已知各个数据的因子算出后，把相对于的已知竞拍价一同输出到BP神经网络中，其中2个因子为输入矩阵，而竞拍价为目标值，通过训练，可以得到一个相应的神经网络，最后利用该神经网络对已知数据进行验证，大概的流程如下：



图8.2 利用BP神经网络进行拟合

进行验证，可得到2者的曲线图和误差分析图为：

图片包含 文字, 地图

描述已自动生成

图8.3 BP神经网络对竞拍价的预测图

图片包含 文字, 地图

描述已自动生成

图8.4 BP神经网络对竞拍价的预测误差图

### 8.2.4差分进化算法

差分进化算法是一种随机的启发式搜索算法，简单易用，有较强的鲁棒性和全局寻优能力。它从数学角度看是一种随机搜索算法，从工程角度看是一种自适应的i的带寻优过程。除了具有较好的收敛性外，差分进化算法非常易于理解与执行，它只包含不多的几个控制参数，并且在整个迭代过程中，这些参数的值可以保持不变。

差分进化算法是一种自组织最小化方法，用户只需很少的输入。它的关键思想与传统进化方法不同：传统方法是用预先确定的概率分布函数决定向量扰动量；而差分进化算法的自组织程序利用种群中两个随机选择的不同向量来干扰一个现有向量，种群中的每一个向量都要进行干扰。



图8.5 差分遗传算法示意图

下面利用差分进化算法计算极大收益：在前面，我们以及利用多元回归法得到了一个计算竞拍成功率的式子，当大于阈值Ps时，有97.7%的概率是竞拍成功的，此时我们认为竞拍成功。由于当竞拍失败时，电视台会出现空白的广告位，就会出现亏损，所以为了极大化电视台的利益，电视台应该保证竞拍的成功。所以要求：



其中广告商的预算、预期销售额和收视率（这里是指以前的收视率）是已知的，而定价是电视台自己根据情况定的，所以可得到式子有：



当保证上面式子成立的情况下，当下面的式子达到极大时，那么电视台的收益就是最大的情况：



其中由于宣传费和底价的定价是电视台自己决定的，可根据具体情况自己决定，所以为自变量，商家个数和宣传费用有比较密切的关系，所以可以计算近似得到。故最终在差分进化算法中输入变量为底价和宣传费用。

# 九、问题五的解

首先，根据问题二中的计算公式，可以计算得到每个电视台各个时间段的收视率。

收视率=P1\*时段活跃指数+P2\*电视台活跃指数+P3\*节目内容（收视人群基数）+P4\*节目收视人群经济实力（P1、P2、P3、P4为不同的权重，在模拟时段：P1=0.5；P2=0.25；P3=0.15；P4=0.1）

收视率=0.5\*时间活跃指数+0.25\*电视台活跃指数+0.15\*节目内容（收视人群基数）+0.1\*节目收视人群经济实力。

这里做出假设，令：

时间活跃指数=1.5

电视台活跃指数=2

节目内容（收视人群基数）=1.5

节目收视人群经济实力=6

可计算得到收视率为2.075；

然后在第四个问题中，我们已经计算出竞拍成功的阈值，以及最高的底价为：



通过分析前面我们建立的数学模型，可以看出收视率、预期销售额、底价以及广告商预算具有很大的相关性，这代表着在试验时如果任意给出三者而不考虑其中的隐含关系，很容易得到与实际情况不符的结果。因此，我们基于问题二的收视率底价模型，并考虑实际情况对上述四者的关系加一定的限制。

①收视率与底价之间基本满足问题二的解；

②平均广告商预算应当在底价附近浮动（多个广告商参与竞价，因此采用平均值更能说明问题）；

③平均预期销售额应当在广告商预算附近浮动（多个广告商参与竞价，因此采用平均值更能说明问题）；

在上述限制的基础上，可以得到更加可靠的数据结果。

其中收视率已经计算得到，而平均广告商的预算、平均预期销售额是未知的。故做进一步的假设为：

平均广告商的预算

平均预期销售额

则通过差分进化算法可得：

收益1

图9.1 差分进化迭代图

最大收益是：23.5466；最优底价、最优所需宣传费是：21.2065、0.5178，其中单位为（1.0万元）。所以最终可得到电视台拍卖该时间段的最大收益为23.5466万元，其中所需要准备的底价为21.2065万元，宣传费为0.5178万元。

进行局部灵敏度分析，根据竞拍时当前市场情况，平均广告商的预算和平均预期销售额会产生相应的波动，所以对该2个变量对最优解和最优值的影响进行研究。

（1）首先，当平均广告商的预算由于市场情况乐观时，提高了10%的预算，即：





而平均预期销售额保持不变，则将参数输入可得到最优：

收益2

最大收益是：25.6625；最优底价、最优所需宣传费是：23.1370、0.6082，其中单位为（1.0万元）。







可得当广告商的预算每增大1%时，其最优底价和最优宣传费应分别提高0.91%和1.745%，而电视台的收益相应提高0.9%左右。

（2）接下来继续考虑，当市场情况较好，广告商将预期销售额提高了10%， 即：





而广告商的预算保持不变，则将参数输入可得到最优：



最大收益是：23.7518；最优底价、最优所需宣传费是：21.3946、0.5224，其中单位为（1.0万元）。







可得当广告商的预期销售额每增大1%时，其最优底价和最优宣传费应分别提高0.089%和0.09%，而电视台的收益相应提高0.087%左右。

从中可以分析得到，当广告商根据市场情况调节预算时，对电视台的最终收益有较大的影响，而单独的广告商的预期销售额的变化对电视台的最终收益影响并不是特别大，但往往当广告商的预期销售额与广告商的预算是相关的。此外，从上面分析还能得到，当市场情况较好时，电视台应该提高相应的宣传费用和底价来增大收益，其中宣传费提升的百分比要比底价高一些。

最后，当竞拍结束后，电视台在最大的收益情况下，应该根据各个广告商的类型依据第3题中的匹配推送表向用户进行推送，即可在极大化电视台收益的情况下提高收视率和销售量。

# 参考文献

1.包子阳，于继周，杨杉. 《智能优化算法及其MATLAB实例》（第2版）.

2. Frank R.Giordano. 《A First course in Mathematical Modeling》（Fifth Edition）.

3.刘毛毛,葛锐,陈钟杰.基于层次分析法的公共检测服务满意度测评建模与实证研究[J].工业计量,2018,28(03):104-106+116.

4.彭道刚,梅兰,李生根,何钧.基于大数据和神经网络的锅炉燃烧含氧量建模研究[J].热能动力工程,2018,33(09):86-92.

5.范柄尧. 差分进化算法在多智能体路径规划中的应用研究[D].太原科技大学,2018.

6.覃强,董建辉.基于多元回归模型的斜坡影响因子分析[J].探矿工程(岩土钻掘工程),2018,45(08):102-106.

# 附录

**程序代码：**

**% 该程序为问题一的层次分析法**

clc;

clear;

close all;

A=[1 1/5 1/7 1/6 1/5;

5 1 1/2 1 1;

7 2 1 1 1

6 1 1 1 1

5 1 1 1 1];

% 平均随机一致性指标由已知资料搜索给出

RI=[0 0 0.52 0.89 1.12 1.26 1.36 1.41 1.46 1.49 1.52 1.54 1.56 1.58 1.59];

n=length(A);

SumA=zeros(n,1);

for i=1:1:n

SumA(i)=1;

for j=1:1:n

SumA(i)= SumA(i)\*A(i,j);

end

end

SumA=SumA.^(1.0/3);

SumA=SumA/(sum(SumA));

W1=A\*SumA;

W=W1/sum(W1);

[~,H]=eig(A);

Hm=0;

for i=1:1:n

Hm=max(Hm,H(i,i)); %求最大特征值

end

V=Hm;

CI=(V-n)/(n-1);

CR=CI/RI(n);

disp('W =');disp(W)

disp('CR =');disp(CR)

if CR<=0.1

disp('一致性可接受');

else

disp('一致性不可接受');

end

**% 该程序为问题三的神经网络算法**

clc;

clear;

close all;

load data2.mat;

pt=result;

%---------------输入数据矩阵---------------%

p=pt(:,1:12).';

%----------------目标数据矩阵--------------

t=pt(:,13:22).';

%------------原始样本归一化-------------%

[P,minp,maxp,T,mint,maxt]=premnmx(p,t);

%------------创建一个新的前向神经网络 ------------%

net=newff(minmax(P),[20,10],{'tansig','purelin'},'traingdx');

%------------设置训练参数--------------------%

net.trainFcn='trainrp';

net.trainParam.show=50;

net.trainParam.lr=0.01;

net.trainParam.epochs=1000;

net.trainParam.goal=1e-3 ;

%------------调用traingdm算法训练BP网络-------------------%

[net,tr]=train(net,P,T);

%--------------对BP网络进行仿真----------------%

A=sim(net,P);

a=postmnmx(A,mint,maxt);

%-------------优化后输入层全职和阈值-------------%

% inputWeights = net.IW{1,1};

% inputbias = net.b{1};

%-------------优化后网络层权值和阈值--------------%

% layerWights=net.LW{2,1};

% layerbias=net.b{2};

%-------------------------------------

x=zeros(12,100);

%-----------------单个广告类型时的向量----------------%

%k代表的是第k类的广告，广告的排列顺序和列表中的顺序相同

k=10;

h=zeros(12,1);

h(1:2)=1;

for i=1:10

if i==k

h(i+2)=1;

end

end

for i=1:1:100

x1(:,i)=[0.01\*i;0;1;1;1;1;1;1;1;1;1;1];

end

h=repmat(h,1,100);

x1=x1.\*h;

x2=x1;

x2(2,:)= 1;

x=[x1 x2];

%--------------计算预测值------------%

x=tramnmx(x,minp,maxp);

Ax=sim(net,x);

%--------------反归一化------------------%

Ax=postmnmx(Ax,mint,maxt);

x=postmnmx(x,minp,maxp);

%-------------将结果归一化到匹配度上-------------%

y=mapminmax(Ax,0,1);

Y=y(k,:);

X=x(1,:);

X=mapminmax(X,19,65);

plot(X(1:100),Y(1,1:100),'b',X(101:200),Y(1,101:200),'r--');

xlabel('年龄(岁)');

ylabel('匹配度');

axis([19 65 0 1]);

legend('男','女');

% -----计算各个年龄段的销售情况 start-----%

for i=1:1:100

if X(i) <= 25

k(1)=i;

elseif X(i) <= 30

k(2)=i;

elseif X(i) <= 45

k(3)=i;

end

end

k1(1)=sum(Y(1:k(1)))/k(1);

k1(2)=sum(Y(k(1)+1:k(2)))/(k(2)-k(1));

k1(3)=sum(Y(k(2)+1:k(3)))/(k(3)-k(2));

k1(4)=sum(Y(k(3)+1:100))/(100-k(3));

disp('(男性)从小到大各个年龄段的销售情况为:');

disp(k1);

k2(1)=sum(Y(101:100+k(1)))/k(1);

k2(2)=sum(Y(101+k(1):100+k(3)))/(k(3)-k(1));

k2(3)=sum(Y(101+k(3):200))/(100-k(3));

disp('(女性)从小到大各个年龄段的销售情况为:');

disp(k2);

% -----计算各个年龄段的销售情况 start-----%

**% 该程序为问题4的神经网络算法**

clc;

clear;

close all;

load data4.mat ns ms X sale bot goal y

f1=0.004\*ns+0.004\*ms+0.242\*X+0.269\*sale+0.271\*bot+0.270\*goal;

f2=0.505\*ns+0.505\*ms-0.004\*X-0.004\*sale-0.005\*bot-0.003\*goal;

H=[f1 f2 y];

save data5.mat H

%%

pt=H;

%---------------输入数据矩阵---------------%

p=pt(:,1:2).';

%----------------目标数据矩阵--------------

t=pt(:,3).';

%------------原始样本归一化-------------%

[P,minp,maxp,T,mint,maxt]=premnmx(p,t);

%------------创建一个新的前向神经网络 ------------%

net=newff(minmax(P),[8,1],{'tansig','purelin'},'traingdx');

%------------设置训练参数--------------------%

net.trainFcn='trainrp';

net.trainParam.show=50;

net.trainParam.lr=0.01;

net.trainParam.epochs=2000;

net.trainParam.goal=1e-3 ;

%------------调用traingdm算法训练BP网络-------------------%

[net,tr]=train(net,P,T);

%--------------对BP网络进行仿真----------------%

A=sim(net,P);

a=postmnmx(A,mint,maxt);

%-------------------------------------

%-----------------验证测试----------------%

x=p(:,4); %输入格式为每个个体为一个列向量

% %--------------计算预测值------------%

x=tramnmx(x,minp,maxp);

Ax=sim(net,x);

% %--------------反归一化------------------%

Ax=postmnmx(Ax,mint,maxt);

x=postmnmx(x,minp,maxp);

Ax

save a\_net net minp maxp mint maxt

**% 该算法为问题4中的差分进化算法**

%%

% 赋初值

clc;

clear;

close all;

warning off;

NP=100; %群体大小

n=2; %个体基因数

F0=0.5; %变异算子

s1\_a=1.0; %对于市场分析可知;

s5\_a=8;

s5\_b=200;

n\_iter=500; %迭代次数

error\_y=1.d-5; %容忍误差值

number=0; %计数器

% 广告商的预算s2,s3为预期销售额,s4为收视率，s5为商家个数,s1为底价

s2=20.85;

s3=27.94;

s4=2.075;

Ps=0.51; %判断阈值

s1\_b=(0.025\*s2+0.002\*s3+0.007\*s4+0.496-Ps)/0.027; %计算最大底价

% 初始化取值

%%

global net minp maxp mint maxt;

load a\_net.mat net minp maxp mint maxt; %读取神经网络

x(1,:)=rand(1,NP)\*(s1\_b-s1\_a)+s1\_a; %随机初始化，每一列代表一个个体，行数=基因数，列数=个体数

x(2,:)=rand(1,NP)\*(s5\_b-s5\_a)+s5\_a;

v=rand(n,NP);

u=rand(n,NP);

u=x;

y1=fun(x);

[temp,y\_pos]=sort(y1); %记录最优历史位置

ymin(1)=temp(1); %记录历史最优值

y\_x=x(:,y\_pos(1));

n\_best=1; %第n个最优值

mseq(1)=0;

%%

for i=1:1:n\_iter

% 变异操作 start

for k=1:1:NP

% 选择三个不同的个体，进行差分 start

r=randperm(NP,4);

r1=r(1);

r2=r(2);

r3=r(3);

if r1==k

r1=r(4);

elseif r2==k

r2=r(4);

elseif r3==k

r3=r(4);

end

% 选择三个不同的个体，进行差分 end

% 计算自适应算子 start

lamb=exp(1-n\_iter/(n\_iter-i+1));

F=F0\*2^(lamb);

% 计算自适应算子 end

v(:,k)=x(:,r(1))+F\*( x(:,r(2)) - x(:,r(3)) );

end

% 变异操作 end

% 进行交叉操作 start

CR=0.5\*(1+rand);

Ln=randi([1,n]);

for k=1:1:n

if rand <= CR || k==Ln

u(k,:)=v(k,:);

else

u(k,:)=x(k,:);

end

end

% 进行交叉操作 end

% 进行边界处理 start

for j=1:1:NP

if u(1,j)>s1\_b

u(1,j)=s1\_b;

elseif u(1,j)<s1\_a

u(1,j)=s1\_a;

end

if u(2,j)>s5\_b

u(2,j)=s5\_b;

elseif u(2,j)<s5\_a

u(2,j)=s5\_a;

end

end

% 进行边界处理 end

% 进行选择操作 start

y=fun(u);

for k=1:1:NP

if y(k) < y1(k)

x(:,k) = u(:,k);

end

end

% 进行选择操作 end

% x是经过选择复制的新群体,u是上次迭代中的旧群体

% 计算适应度，并保留最佳值 start

y1=fun(x);

[temp,y\_pos]=sort(y1); %记录最优历史位置

if(temp(1)<ymin(n\_best))

temp(2)=ymin(n\_best);

n\_best=n\_best+1;

mseq(n\_best)=i;

ymin(n\_best)=temp(1); %记录历史最优值

y\_x(:,n\_best)=x(:,y\_pos(1)); %记录历史最优解

end

% 计算适应度，并保留最佳值 end

if abs(temp(1)-temp(2))<=error\_y

number=number+1;

if number>=5

break;

end

else

number=0;

end

end

%%

% 画图

i

if i>=n\_iter

disp('未达到精度要求');

else

disp('达到精度要求并退出循环');

end

plot(mseq(1:end),50-ymin(1:end))

xlabel('迭代次数');

ylabel('最大收益f(x)');

title('差分进化算法');

str=strcat('最小值是',num2str(ymin(n\_best)));

sk=strcat('最小值点在第',num2str(mseq(end)-1),'次迭代');

text(1,3,sk,'FontSize',12);

text(1,1,str,'FontSize',12);

disp('最大收益是：')

disp(50-ymin(n\_best))

disp('最优底价、最优所需宣传费是')

disp(y\_x(1,n\_best).')

disp((y\_x(2,n\_best)-8).^2\*0.001);

y\_x(2,n\_best)

disp('最小值的迭代次数是');

disp(mseq(end)-1)

%%

% 目标函数

function result=fun(x)

% 广告商的预算s2,s3为预期销售额,s4为收视率，s5为商家个数,s1为底价

global net minp maxp mint maxt;

[~,n]=size(x);

Ax=zeros(1,n);

for i=1:1:n

mx=tramnmx(x(1,i),minp,maxp);

Bx=sim(net,mx);

Ax(i)=postmnmx(Bx,mint,maxt);

end

result=( Ax.\*(1+x(2,:)/500) ) -( x(2,:)-8 ).^2\*0.001;

result = 50-result;

end