**Youth游览疏导及酒店入住规划模型**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 姓名 | 学号 | 学院 | 专业 | 联系方式 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**摘要：**

Youth游乐园即将盛大开园，作为本市建有最多过山车的游乐园，预计届时园区将迎来每天1万人的大客流。针对游乐园的客流问题及附近皇冠假日酒店的房间分配问题，本文分别建立了无序多服务台模型、层次分析模型及ARIMA模型，较为圆满地解决了上述问题。  
 问题一，假定游客在Youth游乐园遵从M/M/n排队论，基于游乐设施承载能力及游览路程、排队时间等，构建了无序多服务台模型。进一步，采用K-约束的贪心算法，得到十条推荐路线（见4.3）。随后，利用层次分析法对游乐设施的游乐设施拥挤程度进行评价，进而对不按照推荐路线的游客及时疏导。  
 问题二，利用附件中的数据，得到酒店每天的房间入住数，并利用ADF方法对其平稳性进行检验，发现该序列是一阶单整的。同时，利用自相关检验和偏自相关检验对模型形式进行选择，从而建立了ARIMA(5,1,5)模型，进而对酒店的入住人数进行预测。特别地，本文还依据AIC和SC准则对模型的滞后阶数进行了检验，验证了模型的合理性。进一步，引入季节、节假日等系列虚拟变量对上述模型进行修正，以得到更为精确的预测结果。在对模型误差项的适应性检验后，发现模型的拟合效果不错。最后，对2016年1月到3月的酒店入住数进行了外推预测。  
  
  
**关键词**：M/M/n排队论；无序多服务台模型；最优路线；ARIMA；适应性检验

1. **序言**

**1.1研究意义**

随着经济的高速发展、人民生活水平逐渐提高，人们对于娱乐生活的要求也越来越高，但同时，人口的持续增长使得许多的娱乐项目总是供不应求。游乐园作为越来越普及化的娱乐设施，过饱和的人流量、游乐设施的超长时间排队经常发生。过长的排队队伍不仅影响游客的游玩体验，并且还会减弱游乐园的人流量疏散，里面的游客没有得到游玩出不去，外面的游客源源不断地进入，游玩体验越来越差，甚至更容易引发一些危险事故，各游乐园都在寻求解决这一现象的方法和途径。如果能够有效地提升游乐园的服务，可以使很多游客得到更好的游玩体验，同时也能提升游乐园的品牌形象。游乐园的游乐设施排队问题和酒店入住问题是影响游乐园游玩体验的重要因素，同时也与游乐园的经济效益息息相关。Youth游乐园作为本市建有最多过山车的游乐园，将在开园后迎来10000人的日平均客流，那么如何疏通客流减少排队成为了游乐园必须解决的问题。而酒店的运营成本较大，如果能够准确地预测出入住酒店的人数，进行适当的运营调整，将能为酒店带来更大的效益。

**1.2问题重述**

针对问题一，Youth乐园有A-J共10个项目点，项目点之间的线路只可按给定路线往返。要保障每位游客都能体验游乐设施的同时，寻找对正在排队的游客进行游览疏导的方法，以此降低游客的排队时间，提高游玩体验。游乐园排队的原因有很多，例如：游乐园内的总人数过多所有设施拥挤；某一游乐设施为该游乐园的招牌项目，极其受欢迎大家都排队；游乐园实际可以承载已入园的游客然而游客们集中在某些设施排队不知道其他设施的空闲程度等。从实质上看，游乐园的排队现象来源于游客们的游玩路线集中而导致的持续性拥堵。因此可以通过一定方法分散游客，达到提升游玩体验的效果。

针对问题二，皇冠假日酒店作为Youth乐园内的酒店，为有需要的游客提供了便利的住所，保障了Youth乐园多日游玩这一模式的成立。已知皇冠假日酒店2015年全年的预订数据，对所给数据进行处理，拟合2015年中的季节、周末等参数对房间预订的影响，运用电脑软件处理数据最终预测2016年1-3月每天预订房间数。

**2、预备知识及符号说明**

**2.1泊松分布**

泊松分布是一种统计与概率学里常见到的离散机率分布，概率函数为

P（X=k）=，k=0,1,2……

其中λ（>0）为常数，则称X服从参数λ的泊松分布，其期望平均值为EX=λ，方差为DX=λ。

**2.2排队论**

排队论是研究系统随机聚散现象和随机服务系统工作过程的数学理论和方法，又称随机服务系统理论，为运筹学的一个分支。一个排队系统由输入过程、排队规则、服务规则、输出过程四个要素组成。输入过程是对顾客到达系统的一种描述，可以是独立的也可以是相关的。排队规则损失制、等待制、混合制。服务规则可以是无窗口、单窗口、多服务窗口。服务规则有先到先服务，后到先服务等，服务窗的服务时间为确定型或者随机型。输出过程是指顾客接受服务完之后离开的过程。

**2.3 M/M/n排队论**

M/M/n排队论用模型表述，令X（t）表示t时刻时，系统中的总顾客数，{X（t），t∈T}是状态空间E={0,1,2，…}上的生灭过程。

生率为：，k=0,1,2，...；灭率为： ，

令，ρ=

，k∈E={0,1,2，…},则当ρ<1时，该生灭过程有平稳分布

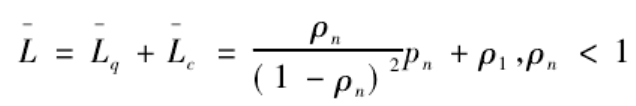
其中， 这里p0表示服务系统空闲的概率。ρ<1为该生灭过程存在平稳分布的充要条件，ρ为系统的服务强度。

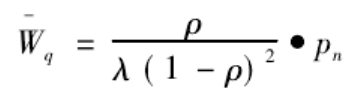
系统中存在n个服务台的时候，顾客需要等待的概率为

其中，

系统内排队等候的顾客数均值，即平均等待队长：

系统平衡时，正在接受服务的顾客数均值为：

L为正在接受服务的顾客数与排队等待服务的顾客和，即队长，为平均队长，其值为

游客排队等候服务的时间均值即平均等待时间为

**2.3 贪心算法**

贪心算法就是在不回溯的前提下寻找解决问题的启发式，在算法执行的过程中希望能够通过每个阶段的局部最优解的选择从而达到全局的最优解。因此能够使用贪心算法的问题应该具有一下两种性质：一是整体的最优解可以通过局部的最优解来逐步得出，而且整体的问题可以划分为多个局部问题，并且对于这些局部问题都能找出各自的最优解。在许多问题的求解中，贪心策略不一定能够产生一个全局最优解，但不管怎样，一个贪心算法的启发式能够在限定的时间内通过局部的最优来近似于全局最优。

贪心算法具有以下元素；

1.一个可以产生解决方案的候选集

2.一个选择函数，通过这个函数可以选择出最佳的候选元素加入到解集中；

3.一个可行性函数，通过可行性函数来决定一个候选元素是否能对解决问题有所帮助；

4.一个目标函数，通过目标函数为问题或者问题的部分指定一个值；

5.一个功能函数，通过功能函数来确定是否得到了一个完整的问题求解。

贪心算法在解决一些数学问题上能够产生很好的解决方案，但是在一些问题上并不能得到最优的解，在用贪心算法解决的许多问题中，有两个很重要的特征；一是贪心选择特点：在处理一个子问题的时候，我们对于这个子问题可以作出任何看起来是最好的选择。由贪心算法作出的选择只是当前子问题的最好选择，并不是整体问题求解的最好选择。在一个选择结束之后重复的进行贪心选择，产生一个个子问题的最优解。换句话说，贪心算法没有回溯，不会当前做出的选择进行重新考虑，这是贪心算法和动态规划算法最主要的不同。动态规划算法在每一步中都会考虑之前的步骤，并且重新考虑之前的算法路径。二是最优子结构：如果问题的最优解包含了每个子问题的最优解。那么我们认为这个问题具有最优的子结构。

**2.4层次分析法**

层次分析法是将与决策有关的影响因素分解为目标、准则、方案等层次，在此基础之上进行定性和定量分析的决策方法。建立递阶层次结构模型，构造出每个层次中的每个判断矩阵，最后计算一致性指标，进行检验。

**2.5 ARMA模型**

ARMA模型即自回归滑动平均模型，是研究时间序列的重要方法。是由AR模型（自回归模型）与MA模型（滑动平均模型）为基础叠加而成。通过对ARMA模型的分析研究，能够更加深刻地理解时间序列的结构与特征，达到最优的预测。

ARMA模型表达式：

注1；Y是预测对象的观测值；

注2；β为自回归系数，α为移动平均系数，都是模型的待估参数；

注3：引入滞后算子B，原模型可以简化为 β(B)Yt＝α(B)

注4：平稳条件是滞后多项式β(B)的根都在单位圆外；可逆条件是滞后多项式α(B)的根都在单位圆外。

**3、排队问题分析**

**3.1游乐场设施排队现象分析**

（图1为游乐设施排队问题的运行图）

**离开**

**游乐设施**

**排队游客**

图1

排队游客有三种机制。1、当设施正在服务时不等待直接离开。2、当设施正在服务时排队等待，不论队伍长度等待时间。3、当设施正在服务时排队等待，但是如果等待的时间预计或者人数长度超过一定数字则离开。

结合实际情况，游乐园中的设施服务流程不考虑工作人员的强行疏导即只考虑前两种排队机制，而一般情况游客会愿意排适合长度的队伍来游玩。因此设施服务流程示意图应为图2所示

空闲

Y

游玩离开

进入设施

到达

Y

N N

忍耐极限

排队结束

排队等待

队伍短

游客

Y N

中途离开

受阻离开

N Y

排队队首

Y N

图2

游客在入园后准备游玩设施等待时的一般心情变化为激动-兴奋-开心-没有感觉-无聊-不耐烦-烦躁-后悔-恼火。在游客情绪下降至一定等级，游客到达忍耐极限，可能使游客选择离开，满意度降低。然而游客在每一个设施的排队情绪以时间来看是总体一致的，那么若能使游客在无法大幅度减少所有设施前总排队时间的前提下，使得游客在每个设施的排队时间尽可能差不多，以达到不会有哪个队伍等待时间过长而不耐烦，每一个队伍都长度都在游客的忍耐极限内的话，游客的体验达到最优。

Youth的预计每日客流量为1万人，游乐设施有10个，承载能力较好。因此可以通过合理安排，使得游客的等待时间在总体上没有大幅改变，而每个游乐设施的等待时间相对平均，不让某个游乐设施等待时间过久导致游客有不良体验，总体使得游客的游园体验情绪没有不满。

**3.2排队算法分析**

游客的一般分布满足泊松分布。游客在游乐园设施排队等待这一事件与M/M/n排队模型契合。游客到达每一个游乐设施的时间间隔服从参数λ的指数分布，当游客到达每一个游乐设施时，若游乐设施没有排队，那么游客可以立即游玩，即不存在排队。若游乐设施正在运行，则游客可以排队等待，直到游玩设施。排队规则采用先到先玩，设施游览的时间服从参数μ的负指数分布。

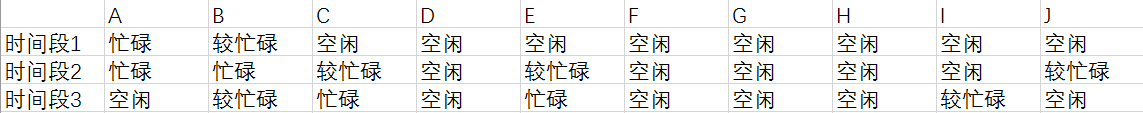
通过对于大量的各类游乐园的一定比较可以发现，游客一般进入游乐园之后都会从距离最近的游乐设施开始体验，并且按照游乐园的地图前往游乐园的下一个最近游乐设施。因此游客自主的游玩过程可以看成是顺序单服务台排队，即：排队-游玩-排队-游玩-排队……直到体验过所有的游乐设施，图3为一批游客进入Youth之后常见游玩顺序可能导致的情况  


图3

根据M/M/n排队模型，可知

队长：， <1

等待时间：

为了减少各游乐设施的排队队长和平均等待时间，将顺序单设施游玩模式变更为无序全设施游玩模式。将游客从一开始就分流，例如可以用园内游览车等工具先将一些游客载去较远处的D、F、H点，然后再交叉游玩。即同时让所有的设施运行起来服务，参见图4。

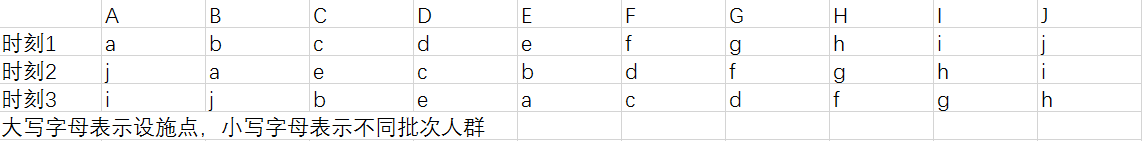


图4

具体的实现可以利用入园口即利用景观车、人力自行车等代步工具分流至不同设施处。将单设施持续服务变为多设施同时服务。这样可以使得游客游玩排队总时间缩短，加快游乐园内的游客流通率，一定程度减少游乐园中等待游玩的游客。

根据附件1，各游乐设施之间的距离是一定的，则可以根据游客的平均步行速度得知游乐设施之间的移动时间也是固定的。于是可以采用多目标线性规划的方式对游玩顺序进行安排规划实现游乐场最大限度接待游客，游客按照建议路线游玩时等待时间较少，每个游乐设施的排队等待时间差距不过大，避免某个设施游客过于集中而造成排队队伍很长的现象。

**3.3基于多约束的k贪心算法**

根据之前的讨论，在实际操作中，游客更多的考虑是在当前游乐设施，去寻找下一个适合的游乐项目，对其他项目不会太过关心。虽然贪心算法能够帮助我们解决路线选择问题，但是贪心算法并不是能够被我们直接应用到游乐园问题的求解中。因为游客每到达一个游乐设施，都会基于这个位置，对其他附近的项目进行重新评分，总体而言是一个动态的过程。所以我们打算基于贪心算法拓展的K-贪心算法来寻找适宜的k条路径，分别对其进行总体评估，最后推荐一条最优的路径给用户。之后在考虑了之前的路径对之后游客的影响后，修正边权矩阵，再迭代十次，求出合适的十条路径。

输入：游客当前所处的游乐设施

输出：从当前所处位置到出口的最优路径

步骤1 初始化，标记每条路径访问过的游乐设施 在图中搜索两点间所有路径，按照深度优先搜索所有可能的从partialPath出发到destination的路径，这些路径中不包含环路

步骤2 根据游客当前所处位置计算出各个其他游乐设施的评分SS，构成修正后的边权矩阵

步骤3 选择前K个评分值最高路径，l表示路径长度，即访问游乐设施的数目

步骤4 判断每条路径是否满足总距离约束，总时间约束，若不满足，则删除该路径

步骤5 对于得到的K个路径，分别对其进行路线多样性的评估，推荐出当前最优的路径

步骤6 改变起点，回到步骤1，进行迭代十次，得出十条推荐路径

**3.3.1游乐设施评分机制**

对于游乐园里的每一个项目，我们的算法在寻找推荐路线之前必须知道，游客在当前位置时，某个游乐设施对游客的吸引程度，即游乐设施对游客的吸引力评分，评分包含三个方面：（1）基于游客到游乐设施的时间的评分，一般来说，距离游客当前位置越远的游乐设施对用户的吸引力相对越低，距离超过900，基本可视作没有吸引力。游客更加愿意选择距离更近的游乐设施。对于距离的评分我们将它定义为DS。（2）基于每个游乐设施每场持续时间的评分，一般来说，每场持续时间越长，游客等待服务的时间越长，吸引力越低，降低游玩体验，我们把它定义为t。（3）基于每个游乐设施每场容纳游客数和每场持续时间的评分，一般来说每场容纳游客数和每场持续时间的比值越大，能够在固定时间内承载更多游客。我们将每场容纳游客数和每场持续时间的比值叫做服务效率CB。我们把DS，CB，α融合起来，形成对每一个游乐设施的评分SS，作为权重，构成边权矩阵，使用MATLAB，进行遍历。



**3.3.2多种评分方式的融合**

在获得每一个游乐设施基于游客到游乐设施时间的评分，每场持续时间的评分和服务效率的评分，进行融合之后得到游乐设施的初步评分，同时引入一个参数α，用于控制基于游客到游乐设施时间的评分，每场持续时间的评分和服务效率的评分的权重。

**3.4小结**

由于早晨开园时游客的进入量是远小于整个游乐园游乐设施的承载量的，因此在最开始给与游客建议路线，使得游客们可以根据自己的喜好，最先玩到最想要玩的设施，同时也使得游客们分散开来，不会拥挤在游乐园门口的设施，以此得到最佳的游园体验。当进入游乐园的游客人数超过一定数量时，所有设施都处于饱和状态，这样的几条路线也可以维持这些设施的排队程度的相对平均，并且给与还想继续在整个游乐园游玩的游客最佳的游玩路线。对于不愿意走建议路线的游客，可以由游乐园方在电子屏实时公告排队人数和等待时间，推荐游玩方向，疏导过长的队伍。

**4、排队模型建立**

**4.1模型假设**

不计一批游客进入设施的时间损失

假设所有设施开园即投入工作，不考虑维修状态

假设没有游客会迷路

假设不会使游客不耐烦，因此以第二种排队机制规划游客的游玩路线。

假设少数游客不按照规划路线游玩没有影响

假设不考虑用餐时间对客流的影响

**4.2数据处理**

**4.2.1**根据每个游乐项目的每场容纳游客数和每场持续时间得到每个游乐项目的承载力参数。根据一定时间（超过33分钟）游乐设施可以服务人数的最大值计算游乐项目的承载力，得到服务效率由小到大排序为D<A<I<G<C=E=F=H<B<J

**4.2.2** 为了保证游乐体验，和游乐园的可持续发展，通过计算游客的休息时间，得出推荐路线包含8个游乐设施。

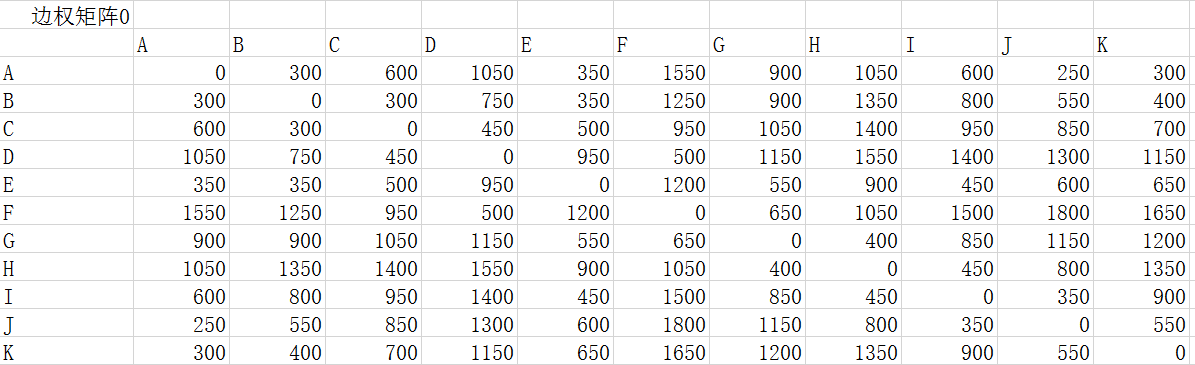
**4.2.3**将附件1的游乐园规划图简化成点线距离图，设出入口为K点。计算出任意两个点之间的距离，非直达（经过别的点）之间的两个点取最短路径距离。（图5）

图5

由于要选取尽量短的路径达到最优体验并且考虑到电脑的计算能力简化计算，超过900米以上距离的路线。设置为无穷大（图6为计算使用的边权矩阵）

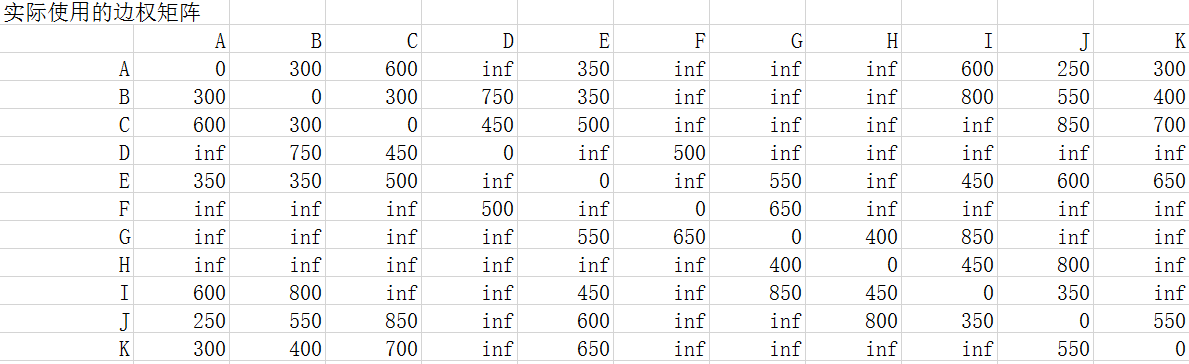


图6

并且得到第一条推荐路线：J-A-B-C-D-F-G-E-K

**4.3多目标规划设计首游乐设施的游览路线**

计算从十个游乐设施分别出发，总共规划十条路线，始目标依据服务效率排序从大到小计算，顺序为J-B-H-F-E-C-G-I-A-D



以B为起点得到推荐路线B-C-E-G-H-I-J-A-K



以H为起点得到推荐路线H-G-F-D-C-B-E-A-K



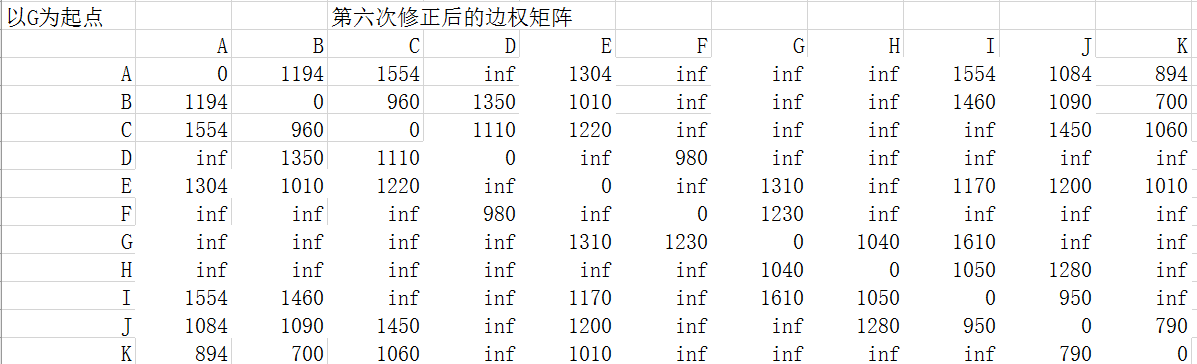
以F为起点得到推荐路线F-D-C-B-E-I-J-A-K



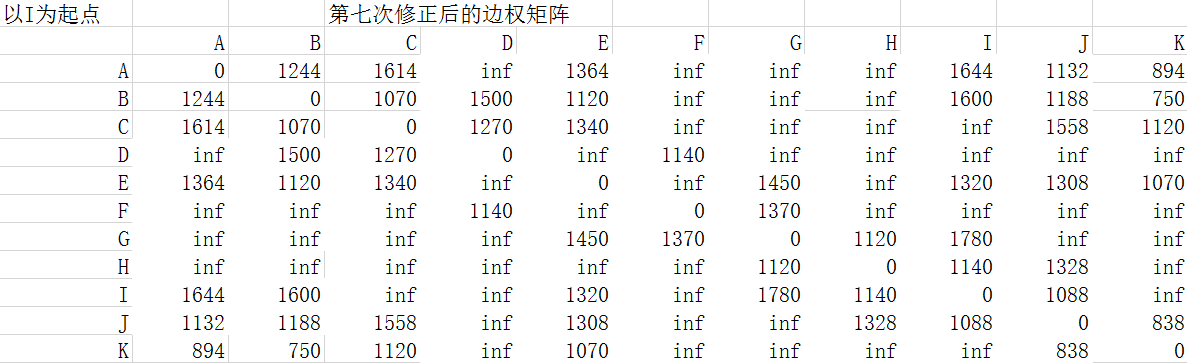
以E为起点得到推荐路线E-G-H-I-J-A-B-C-K



以C为起点得到推荐路线C-B-E-G-H-I-J-A-K



以G为起点得到推荐路线G-F-D-C-B-E-I-J-K



以I为起点得到推荐路线I-H-G-F-D-C-E-J-K



以A为起点得到推荐路线A-J-I-H-G-E-C-B-K



以D为起点得到推荐路线D-F-G-H-I-J-E-B-K

以游乐园中每一个设施为起点规划，共得到十条不同的游玩路线，考虑到游客不同的喜好，可以供游客选择不同的路线。

对于不按照推荐路线游玩的游客，由游乐园方提供电子屏实时显示各设施目前的排队人数，等待时间预计，给与游客自主选择。

**4.4 模型检验**

简单计算项目的适用效率，得图7，知十条规划路线较为合理。

C:\Users\佳琳\Documents\Tencent Files\1515620037\Image\Group\9E43F90@)0_E)D`KC{FK160.png

图7

再通过计算机仿真平台模拟运营，将十条建议路线的系统概念模型转化为计算机仿真模型，并进行模拟分析，对结果分析研究。

（可能的话）在试运营中向游客发放问卷，了解影响游客等待心理变化和游客等待行为，接受等待时间等，适当给与调整方案。

**4.5模型缺点**

1. 推荐路线固定，无法满足所有游客的个性化需求
2. 模型较为死板，不够灵活

**4.6模型的优化**

以上的十条规划路线主要通过分散的无序多服务台服务减轻排队程度，对于游乐园方操作极为方便。在优化中，选择可以更机动的方案作为辅助。

**4.6.1层次分析法**

游园设施疏导指数施的硬件条件和动态的当时设施所容纳的游客人数，通过层次分析法把问题条理化，层次化。将设施的硬件条件分为三个重要的判断因素，分别为游乐设施的（1）承载能力（2）周围有几条路线（3）周围设施的承载能力。再结合当时游园设施里的排队时长和周边游园设施的排队时长，建立递接层次模型，游园设施疏导指数评价因子判断矩阵A

步骤一，构造出准则层中的判断矩阵，引用数字1-9及其倒数作为标度来定义判断矩阵A，B1（见图8）

|  |  |
| --- | --- |
| 标度 | 定义 |
|
| 1 | i因素比j因素同等重要 |
| 3 | i因素比j因素略重要 |
| 5 | i因素比j因素重要 |
| 7 | i因素比j因素重要得多 |
| 9 | I因素比j因素绝对重要 |
| 2,4,6,8 | 介于以上几种判断之间 |

图8

步骤二，进行一致性验证。

先计算一致性指标，



再计算一致性比例

CR，（ RI为平均随机一致性指标）

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 阶数n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| RI | 0 | 0 | 0.52 | 0.89 | 1.12 | 1.24 | 1.36 | 1.41 | 1.46 |

图9

（图9：平均随机一致性指标）

得到CR<0.1，即判断矩阵一致性可以接受。

得到判断矩阵，

A= ，

为了数据的精确，对于层次分析法的权重向量W进行两种方法的计算（具体代码见附录4）

1. 几何平均法（方根法）利用公式将判断矩阵元素按行相乘得到新的向量，再将新的向量的每个分量开n次方，所得向量归一化，得到权重向量。
2. 算数平均法（求和法）将判断矩阵归一化，将归一化后的各列相加，相加后的向量除以n即得权重向量

（A向量的归一化指求）

对游乐设施硬件条件B1采用一致性指标较低的算数平均算法

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 游乐设施硬件条件B1 | | | | | | |
|  | W1 | W2 | W3 | λ max | CI | CR |
| 几何平均法 | 0.6716 | 0.2654 | 0.0629 | 3.02896 | 0.0144 | 0.02497 |
| 算术平均法 | 0.6688 | 0.2673 | 0.0637 | 3.0287 | 0.01439 | 0.02481 |

得到W1=0.6688,W2=0.2673,W3=0.0637,分别为承载能力，周围有几条路线，周围设施的承载能力的权重

对游乐设施疏导指数A采用一致性指标较低的算数平均算法

得到L1=0.6695,L2=0.2516,L3=0.0787,分别为游乐设施硬件条件，游园设施里的排队时长和周边游园设施的排队时长的权重

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 游乐设施疏导指数A | | | | | | |
|  | L1 | L2 | L3 | λ max | CI | CR |
| 几何平均法 | 0.7393 | 0.1998 | 0.0607 | 3.0571 | 0.02855 | 0.04922 |
| 算术平均法 | 0.6695 | 0.2516 | 0.0787 | 3.0568 | 0.0284 | 0.04896 |

由权重表格可以知道，对硬件而言游乐设施的各条件中承载能力最重要，周围几条路线第二，周围设施的承载能力最后；对游乐设施疏导指数而言，硬件条件最重要，游乐设施的排队时长第二，周边游乐设施的排队时长权重最次。

**优化模型1**

假设游客依照推荐路线游玩，又游乐园过饱和，排队实在太长，可以根据层次分析法中得到的游乐设施疏导指数，决定当队伍超过一定长度时，跳过某些设施直接去往下一个游玩体验更优（需要调查游客的想法）可以达到游玩时间固定，考虑不游玩某些设施，使得游玩时间与总等待时间比最大的效果。

**优化模型2**

通过层次分析法得到的各设施的疏导指数，将不按照推荐路线游玩的游客推荐往疏导指数最高的游乐设施游玩，实时给与游客游玩方向的推荐和疏导，以平衡整个游乐园的排队时长。同时还可以解决设备维修状态的通知疏导。

**模型缺陷**

不考虑游乐场的热门项目可能无法疏导

无法给出所有可能的刚好达成游客期望游览项目的路线，只能一定程度提升满意度。

完全的不按照推荐路线游玩的游客无法计入，只能以动态广播或电子屏辅助疏导。

**4.7模型的推广**

路线规划模型可以用于各旅游景点，包括给与旅行社多支队伍同时在一个景区游玩的路线规划模型

动态疏导可以用于其他游乐设施或者大型场馆。

**4.8展望与建议**

1.排队区的条件优化。必须明确排队的公平性，才能保证游客的心理平衡。强化排队区域的服务，如遮阳（雨）棚等，也可在炎热的夏天增设移动贩卖冰水扇子等的工作人员。

快速通道排队法。迪斯尼乐园运行一种名为快速通道的排队方法。运作流程为先让游客去游乐设施旁的机器里拿一张卡片，上面记录了一个时间，在此时间之后可以凭该卡片从快速通道进入该游乐设施。通过系统计算后台控制快速通道的队伍不会过长，并且卡片本身增强了游乐园的娱乐体验。由于需要增设机器增大运营成本，故仅供参考。

2.为了更好地鼓励游客依照推荐路线游玩，可以向愿意按照路线的游客提供人力自行车、游览车等代步工具。

3．可以将十种不同的游玩路线加入故事情节或者制成十种不同的卡片，既有游乐园收藏价值，同时增加了娱乐价值，还可以吸引游客下次再来收集别的路线卡片。

**5、第二问问题分析**

游乐园的酒店设施在一定程度上优化了游乐园的客户体验，可以给与游客更多的游玩选择。然而，酒店的运营成本是相对较大的。酒店的收入源于游客的消费，因此酒店的入住人数的多少对于酒店的运营非常重要。然而游客是否入住酒店有多方面的原因，例如季节、周末、法定假日、寒暑假等。皇冠假日酒店作为游乐园内的酒店，需要为游客提供更为优质的服务，在保证服务到位的前提下，可以通过一定的判断减少运营成本。如果能够建立数学模型，预测之后时间里的酒店入住人数，就能够更加合理的分配资源，提高对游客的服务效率，安排服务人员，最大化利用服务资源。那么如何能够预测之后的入住人数呢?可以从不同的角度入手建立模型。大体来看,可以分为两种类型

**5.1结构模型法**

即针对实际问题中的特定情况，简化问题,根据相关的经济理论,或者从经济模型进行定性分析,找出相关因素作为解释变量,取得有关数据,然后建立回归模型。在具体计算中,可以把酒店入住人数作为被解释变量Y,选取影响酒店入住人数的主要因素作为解释变量例如,季节,工作日/周末,法定假日,暑期等，调查取得有关数据后建立回归模型,估计参数,找出酒店入住人数与有关主要因素的数量关系,然后对有关因素分别进行预测,再代入回归方程,得到酒店入住人数的预测值。

**5.2数据驱动建模法**

把经济变量的观测值看作是由随机数据生成过程生成的,在对该生成过程做出一定的假定之下,开展模型的参数估计及推断工作。具体来说,分析由酒店入住人数的历史数据构成的时间序列的特征,建立时间序列模型,并据此做出预测。

**5.3比较两类方法**

可以看出:第一类方法思路较为自然,方法也相对简单,但是如何确定和取舍影响解释变量的主要因素,即解释变量的选取问题,目前并没有较成熟的经济理论作指导,对于该问题的经济分析因人而异,而且原始数据不易取得,这两个原因会直接影响模型的建立和预测效果。第二类方法,原始数据较容易取得,但由于数据包含了各方面的信息,具有自身的特殊性,如数据可能包含有趋势、季节性、异常值、条件异方差等,使得简单的模型难以充分描述样本数据的特征,因而预测的精度对模型的依赖型较强。   
本文拟采用第二类方法,从数据出发,采用某酒店一段时间的酒店入住人数数据构成的时间序列,在一定的假定下先建立较为简单的模型,然后放宽假定,改进模型,以达到较佳的预测效果。

1. **第二问模型建立**

利用matlab整理预定房间数据（见附录5）得到处理好的数据。

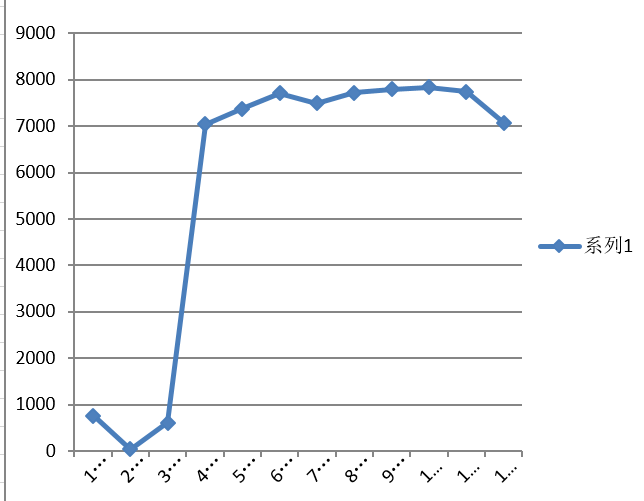


图10  
将数据在excel中简单处理，得到图10（每月预定房间数量折线图），可以看出2月在1.3月之间明显预定房间数少，4月开始数据大幅高于1-3月数据。推断1-3月预定房间数较少，2月可能受春节影响。

采用以2015年1月1日到2016年2月1日的数据作为回归样本，除时间 t、虚拟变量h、v、wy等自定义外, 其它数据来自对题目附件数据的处理。利用EVIEWS计算。

**6.1酒店预定房间数Xt的平稳性检验**



图11

由图11观察得出，该序列显著不平稳。

进一步考察序列x的自相关图和偏相关图

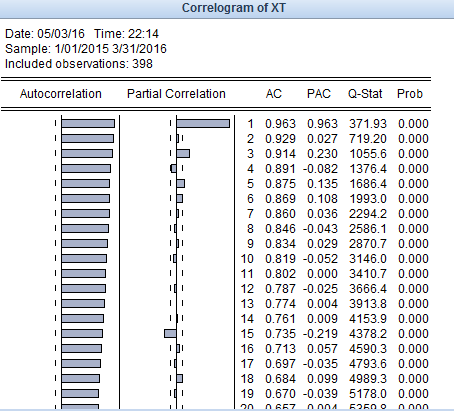


图12

从图12自相关系数可以看出，衰减到零的速度非常慢，所以断定x序列非平稳，为了证实这个结论，进一步对其做ADF检验：

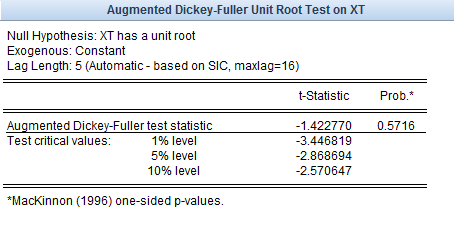


图13

结果见图13可以看出，在显著性水平0.05下，接受存在一个单位根的原假设，进一步验证了序列的不平稳。为了找出其非平稳的阶数，需要对其一阶差分序列等进行ADF检验。

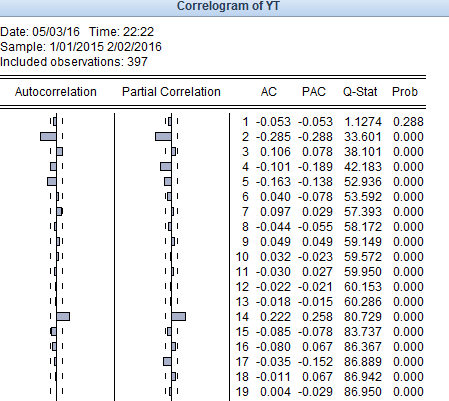
**6.2差分次数d的确定**

X序列显著非平稳，现对其一阶差分进行ADF检验。在对x的一阶差分序列进行ADF检验前，需要明确x的一阶差分序列的趋势特征。其时序图（图14）如下：

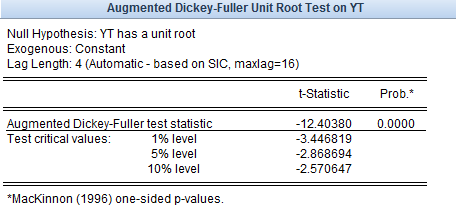


图14

由Xt的一阶差分序列的时序图可见，一阶差分序列不具有趋势特征，且具有零均值。进一步考察序列dx的自相关图和偏相关图：



从自相关系数可以看出，衰减到零的速度非常快，所以断定dx序列平稳，为了证实这个结论，进一步对其做ADF检验：



可以看出在显著性水平0.05下，拒绝存在单位根的原假设，说明序列dx的一阶差分序列是平稳序列，因此**d=1**。

**6.3模型定阶**

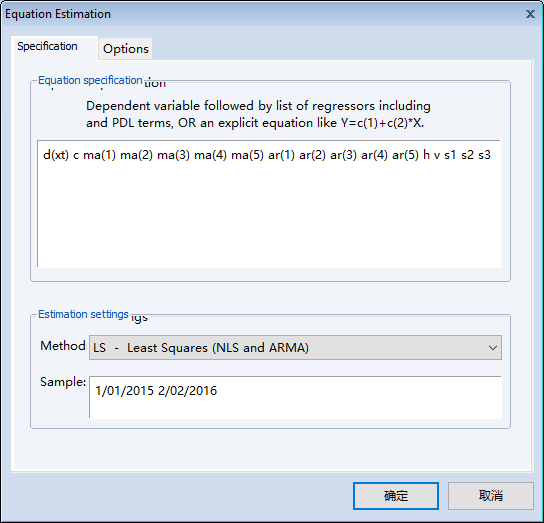
由序列Yt的自相关偏相关图可以看出，自相关和偏相关系数在K=6后很快趋于0，我们初步建立ARMA(5，5)模型。

接下来借助模型定阶方法确定：分别估计AR(1)~AR（6）以及MA(1)~MA(6)模型，并将模型估计结果中的AIC信息准则值记录下来：

由图形可见，AIC准则确定的AR和MA模型阶数均为5，即选择ARMA(5，5)模型。

**6.4模型参数估计**

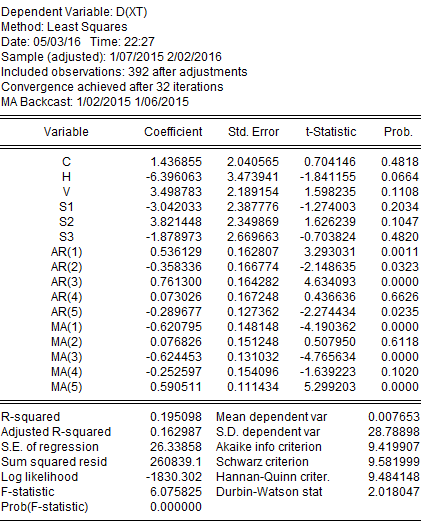
可用菜单或者命令两种方式建立。在主菜单选择Estimate Equation，得到



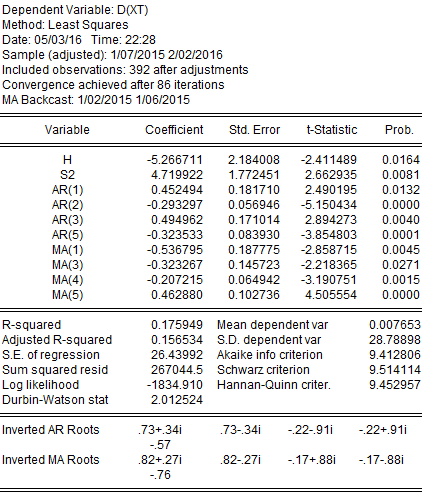
我们加入了可能影响酒店入住人数的因素，包括寒暑假（v）,法定假日及周末（h），季节（s1表示冬天，S2表示春天，s3表示夏天）的虚拟变量。

（取值说明：以v为例，2015/2016年2月15-3月15，2015年7月1日-8月31日 取值为1， 其他为0；s1中属于冬季的日期取值为1，其他日期取值0）

估计结果如下：



由估计结果中相伴概率（Prob.）的值看出，常数项C，v，S1，S3，AR(4),MA（2）系数不显著，故剔除这些项重新进行估计，最后得到如下ARMA(5，5)模型。



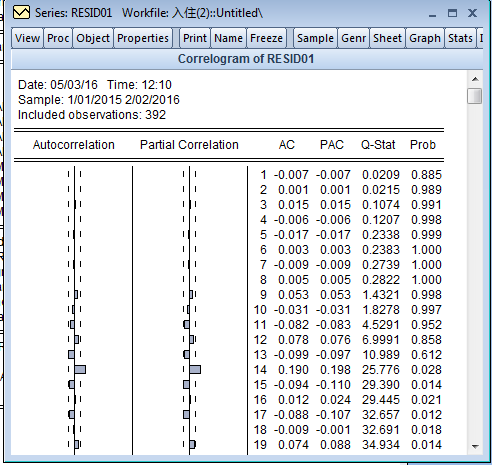
综上可见，我们对同一个平稳序列建立多个适合模型，但比较AIC和SC的值，以及综合考虑其他检验统计量，我们认为ARMA（5,5）是较优选择。但从估计结果中我们可以看到R方的值为0.175拟合效果一般，之后我们将会从虚拟变量方面进行改进，以提高模型的拟合优度。

**6.5模型适应性检验**

参数估计后，应对拟合模型的适应性进行检验，实质是对模型残差序列进行白噪声检验。若残差序列不是白噪声，说明还有一些重要信息没被提取，应重新设定模型。可以对残差进行纯随机性检验，也可用针对残差的检验。

残差序列从01/01/2015至02/02/2016采用拟合的ARMA(5,5)模型生成。

序列的自相关偏自相关图如下：



（偏）相关函数值、以及Q-Stat及其p值显示，残差序列不存在自相关，为白噪声，因此模型是适合的模型。

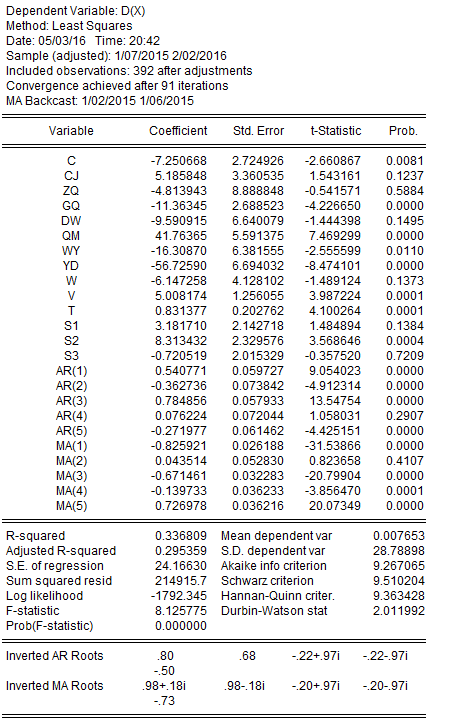
模型拟合图如下：



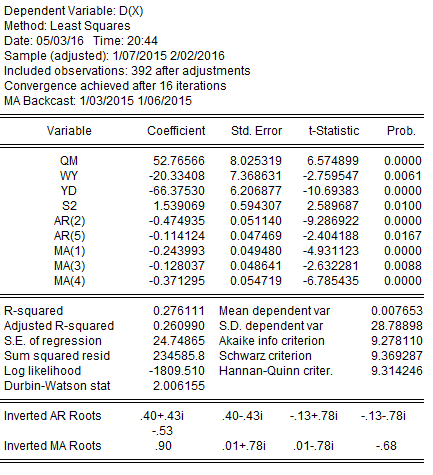
**6.6模型的改进：**

上述模型中我们将法定假日以及周末放在一起作为一个虚拟变量，可能会由于平均效应弱化了某些法定假日的影响（比如元旦，春节等），同时，我们看到季节效应对模型的影响不是很明显。因此我们把每个法定假日单独作为虚拟变量（元旦yd，春节cj，清明qm，五一wy，端午dw，国庆gq，中秋zq）引入模型，同时将月份作为虚拟变量加入模型（t=1,2…依次类推），周末用w表示，寒暑假为v。

（取值说明，例如是元旦的取1，不是取0；其他以此类推）



从估计结果中我们可以看到R方的值为0.336809，拟合效果有了较大的提高。从相伴概率（Prob.）的值看出，春节（cj），中秋节（zq），端午节（dw）,周末（w），冬季（s1）,夏季（s3）,AR(4),MA（2）,系数不显著，故剔除这些项重新进行估计，并重复剔除本步骤，最后得到如下参数估计结果：



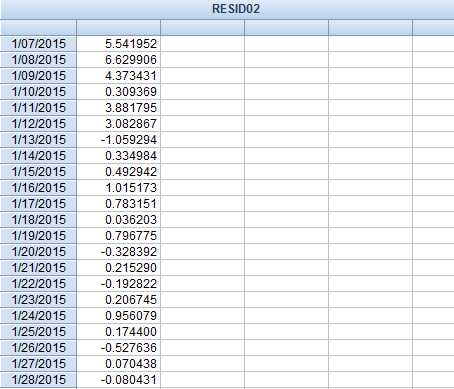
最终得到如下ARMA(5,4)模型：

，AIC=9.278110，SC=9.369278

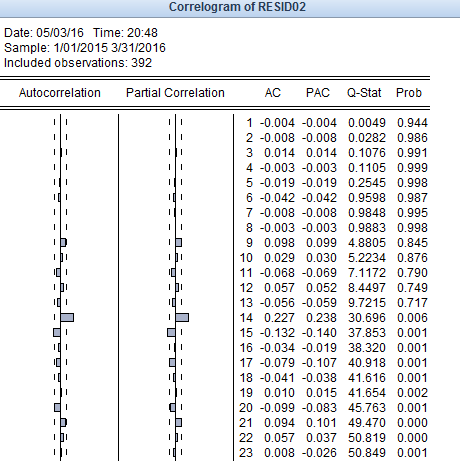
**6.7模型的适应性检验**

参数估计后，应对拟合模型的适应性进行检验，实质是对模型残差序列进行白噪声检验。若残差序列不是白噪声，说明还有一些重要信息没被提取，应重新设定模型。可以对残差进行纯随机性检验，也可用针对残差的检验。

**（1） 残差序列的生成**

残差序列从1/01/2015至2/2/2016采用拟合的ARMA(5,4)模型生成， 得到残差序列resid02 ：

序列的自相关偏自相关图如下：



（偏）相关函数值、以及Q-Stat及其p值显示，残差序列不存在自相关，为白噪声，因此模型是适合的模型。

**6.8模型预测**

预测方法常用有两种：Dynamic forecast和Static forecast：

动态预测是根据所选择的一定的估计区间，进行多步向前预测（**从预测样本的第一期开始计算多步预测**）,每一步都是采用前面的预测值计算新的预测值。而样本范围内（1/01/2015 2/02/2016）的序列实际值是已知的。因此，动态预测只是适应于样本外（2/03/2016-3/31/2016）预测，而不适应于样本内（1/01/2015 2/02/2016）预测。

**静态预测**是滚动的进行向前一步预测，即每预测一次，用**真实值**代替预测值，加入到估计区间，再进行向前一步预测（**利用滞后因变量的实际值而不是预测值计算一步向前（one-step-ahead）预测的结果**）. 对于样本外（2/03/2016-3/31/2016）的预测**需要提供样本外预测期间的解释变量值。对静态预测，还必须提供滞后因变量的数值。而对于样本外**（2/03/2016-3/31/2016）**的预测通常因变量的实际观测值是未知的，所以，静态预测一般只适应于样本内**（1/01/2015 2/02/2016）**预测，不适应于样本外预测（只可以进行向前一步样本外预测）。**

综上所述，在预测时，样本内（1/01/2015 2/02/2016）预测选用静态预测或模型的拟合值（估计值），样本外（2/03/2016-3/31/2016）预测选用动态预测。

**6.8.1具体步骤：**

（1）进行样本内（1/01/2015 2/02/2016）静态预测：

预测序列记为xf\_static, 预测方法选择“Static forecast”，预测样本区间为“1/01/2015 2/02/2016”，点击OK，得到下列预测图



（2）进行样本外（2/03/2016 3/31/2016）动态预测。

预测序列记为xf\_dynamic，预测方法选择“Dynamic forecast”，预测样本区间为样本外区间**“2/03/2016 3/31/2016”**，点击OK，得到下列预测图:

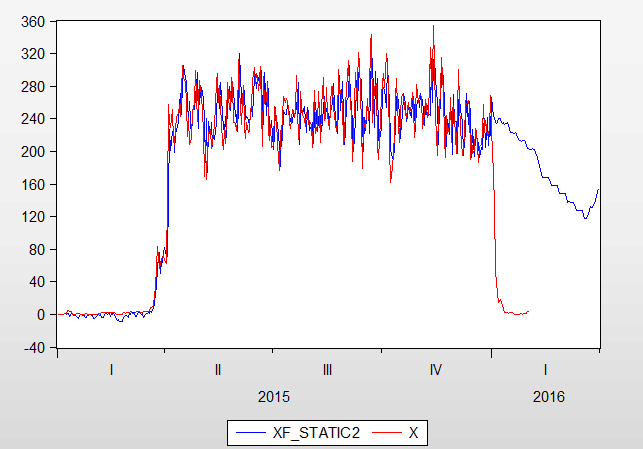


（4）建立新的序列xf\_fit存放序列的静态和动态预测值，在样本内（1/01/2015 2/02/2016）采用静态预测值或序列的拟合值（估计值），在样本外（2/03/2016 3/31/2016）采用动态预测值。所以，将序列xf\_static中1/01/2015到2/02/2016的数值复制到序列xf\_fit的对应位置，然后将序列xf\_dynamic中2/03/2016到3/31/2016的数值复制到序列xf\_fit的对应位置，这样得到的序列xf\_fit就是序列x的预测值。

（5）将序列的实际值x和这里的预测序列xf\_fit以组的形式打开，组序列图形，即x的实际值与预测值图形如下：



作为对比，使用最初的模型进行了预测，得到下图:



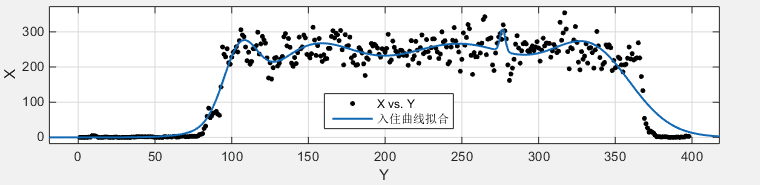
可以看到，改进后的模型预测结果比最初了模型有了很大的改善。

最后，由改进后的模型得到2016年1-3月份每天的预订人数：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 日期 | 预订人数 | 日期 | 预订人数 | 日期 | 预订人数 |
| 2016/1/1 | 194 | 2016/2/1 | 6 | 2016/3/1 | 5 |
| 2016/1/2 | 129 | 2016/2/2 | 4 | 2016/3/2 | 5 |
| 2016/1/3 | 80 | 2016/2/3 | 5 | 2016/3/3 | 5 |
| 2016/1/4 | 92 | 2016/2/4 | 5 | 2016/3/4 | 5 |
| 2016/1/5 | 28 | 2016/2/5 | 5 | 2016/3/5 | 5 |
| 2016/1/6 | 48 | 2016/2/6 | 5 | 2016/3/6 | 5 |
| 2016/1/7 | 23 | 2016/2/7 | 5 | 2016/3/7 | 5 |
| 2016/1/8 | 32 | 2016/2/8 | 5 | 2016/3/8 | 5 |
| 2016/1/9 | 35 | 2016/2/9 | 5 | 2016/3/9 | 5 |
| 2016/1/10 | 29 | 2016/2/10 | 5 | 2016/3/10 | 5 |
| 2016/1/11 | 24 | 2016/2/11 | 5 | 2016/3/11 | 5 |
| 2016/1/12 | 21 | 2016/2/12 | 5 | 2016/3/12 | 5 |
| 2016/1/13 | 19 | 2016/2/13 | 5 | 2016/3/13 | 5 |
| 2016/1/14 | 17 | 2016/2/14 | 5 | 2016/3/14 | 5 |
| 2016/1/15 | 15 | 2016/2/15 | 5 | 2016/3/15 | 5 |
| 2016/1/16 | 14 | 2016/2/16 | 5 | 2016/3/16 | 5 |
| 2016/1/17 | 14 | 2016/2/17 | 5 | 2016/3/17 | 5 |
| 2016/1/18 | 12 | 2016/2/18 | 5 | 2016/3/18 | 5 |
| 2016/1/19 | 11 | 2016/2/19 | 5 | 2016/3/19 | 5 |
| 2016/1/20 | 10 | 2016/2/20 | 5 | 2016/3/20 | 5 |
| 2016/1/21 | 8 | 2016/2/21 | 5 | 2016/3/21 | 5 |
| 2016/1/22 | 7 | 2016/2/22 | 5 | 2016/3/22 | 7 |
| 2016/1/23 | 7 | 2016/2/23 | 5 | 2016/3/23 | 8 |
| 2016/1/24 | 6 | 2016/2/24 | 5 | 2016/3/24 | 10 |
| 2016/1/25 | 6 | 2016/2/25 | 5 | 2016/3/25 | 11 |
| 2016/1/26 | 6 | 2016/2/26 | 5 | 2016/3/26 | 13 |
| 2016/1/27 | 5 | 2016/2/27 | 5 | 2016/3/27 | 14 |
| 2016/1/28 | 4 | 2016/2/28 | 5 | 2016/3/28 | 16 |
| 2016/1/29 | 5 | 2016/2/29 | 5 | 2016/3/29 | 17 |
| 2016/1/30 | 4 |  |  | 2016/3/30 | 19 |
| 2016/1/31 | 4 |  |  | 2016/3/31 | 20 |

**6.8 模型的检验**

用matlab的cftool将数据进行曲线拟合（代码见附录6）



根据拟合的曲线检验ARIMA模型得到的预测数据，基本一致，表示模型可信度。

参考文献

1. 王仁志 苗维亚. 基于M/M/n排队论的大型旅游景区内部排队现象研究. 1006-012X（2012）-03-0177（04）
2. 罗芬 廖薇. 主题公园游乐设施游客等待心理变化研究——以长沙世界之窗为例. 1673-9272（2010）03-0051-04
3. 王保淳. 旅游景区空间布局与服务设施规模规划研究：[硕士学位论文]. 大连：大连理工大学，2014
4. 陆国锋. 基于多约束多目标的旅游路线推荐及关键算法研究：[硕士学位论文]. 长沙：国防科学技术大学，2014
5. 陆国锋 黄晓燕 吕绍和 王晓东. 基于互联网信息的多约束多目标旅游路线推荐. 长沙410073；成都610017
6. 孙辉. 排队长度约束下的瓶颈交叉口协调控制方法研究：[研究生学位论文]. 成都：西南交通大学，2010
7. Eveati. Ant colony optimization-The basic concept and matlab implementation: Machine Learning, 2014
8. 刘中学 张毅. 包含虚拟变量的澳门酒店入住率预测模型及实证. F224.7，1002-6487
9. 陈晓卫 李洪. 酒店入住人数模型研究. China academic journal electronic publishing house. 总第78期
10. 邓雪 李家铭 曾浩健 陈俊羊 赵俊峰. 层次分析法权重计算方法分析及其应用研究. 广东：华南理工大学 510640）
11. 卢芳. 基于排队论的电动汽车充电站选址定容研究：[硕士专业学位论文]. 北京：北京交通大学，2015

**附录1 遍历找出从某一起点到出口的所有路径matlab源代码**

function possiablePaths = findPath(Graph, partialPath, destination, partialWeight)

pathLength = length(partialPath);

lastNode = partialPath(pathLength);

 nextNodes = find(0<Graph(lastNode,:) & Graph(lastNode,:)<inf); GLength = length(Graph);

possiablePaths = [];

if lastNode == destination possiablePaths = partialPath;

 possiablePaths(GLength + 1) = partialWeight;

 return;

elseif length( find( partialPath == destination ) ) ~= 0

 return;

end for i=1:length(nextNodes)

 if destination == nextNodes(i)

 tmpPath = cat(2, partialPath, destination);

 tmpPath(GLength + 1) = partialWeight + Graph(lastNode, destination);

 possiablePaths( length(possiablePaths) + 1 , : ) = tmpPath;  nextNodes(i) = 0;

 elseif length( find( partialPath == nextNodes(i) ) ) ~= 0

 nextNodes(i) = 0;

 end

end

nextNodes = nextNodes(nextNodes ~= 0);

 for i=1:length(nextNodes)

 tmpPath = cat(2, partialPath, nextNodes(i));

 tmpPsbPaths = findPath(Graph, tmpPath, destination, partialWeight + Graph(lastNode, nextNodes(i)));

 possiablePaths = cat(1, possiablePaths, tmpPsbPaths);

end

**附录2 路线1为例的绘图程序代码（坐标来自PS描点）**

%绘制原始图

C=[

20.12 5.62

23.77 6.66

27.77 7.95

28.89 14.75

20.15 10.12

21.47 13.26

12.19 13.67

6.84 11.03

13 7.54

17.17 3.65

21.70 0.03

];

t=0:0.01:1;

N=11;

A=[1 2 3 4 6 7 8 9 10 1 11];

for m=1:N-1

x=t\*C(A(m),1)+(1-t)\*C(A(m+1),1);

y=t\*C(A(m),2)+(1-t)\*C(A(m+1),2);

plot(x,y,'k--')

hold on

end

A=[5 1;5 2;5 3;5 7;5 9];

for m=1:5

x=t\*C(A(m,1),1)+(1-t)\*C(A(m,2),1);

y=t\*C(A(m,1),2)+(1-t)\*C(A(m,2),2);

plot(x,y,'k--')

hold on

end

x=t\*C(2,1)+(1-t)\*C(11,1);

y= t\*C(2,2)+(1-t)\*C(11,2);

plot(x,y,'k--')

hold on

text(C(1,1),C(1,2),'A');hold on

text(C(2,1),C(2,2),'B');hold on;

text(C(3,1),C(3,2),'C');hold on;

text(C(4,1),C(4,2),'D');hold on;

text(C(5,1),C(5,2),'E');hold on;

text(C(6,1),C(6,2),'F');hold on;

text(C(7,1),C(7,2),'G');hold on;

text(C(8,1),C(8,2),'H');hold on;

text(C(9,1),C(9,2),'I');hold on;

text(C(10,1),C(10,2),'J');hold on;

text(C(11,1),C(11,2),'出入口');hold on;

%图一

t=0:0.01:1;

N=9;

E=[ 10 1 2 3 4 6 7 5 11];

for m=1:N-1

x=t\*C(E(m),1)+(1-t)\*C(E(m+1),1);

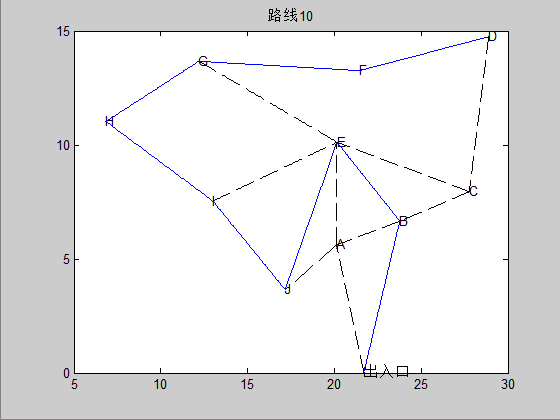
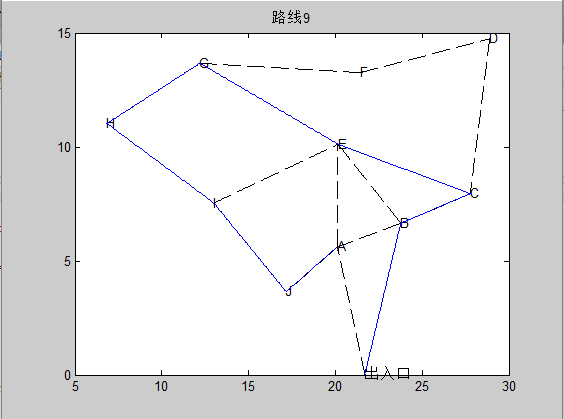
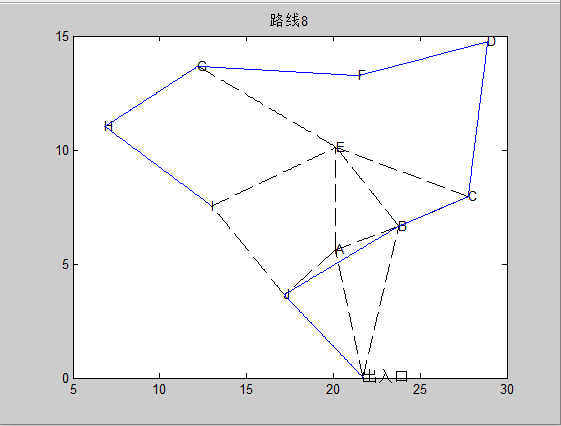
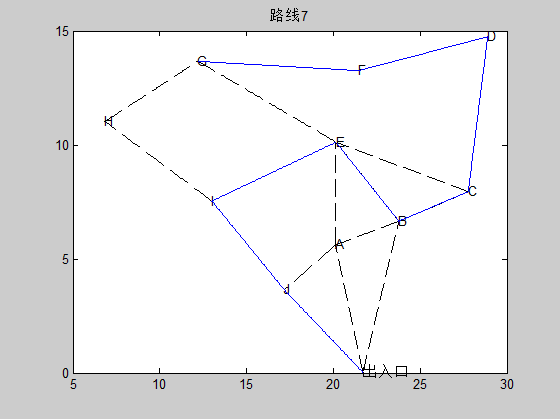
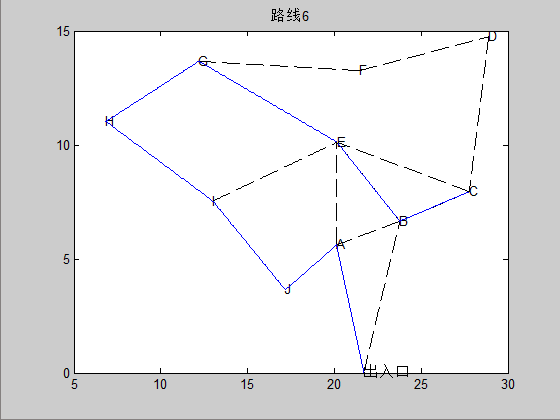
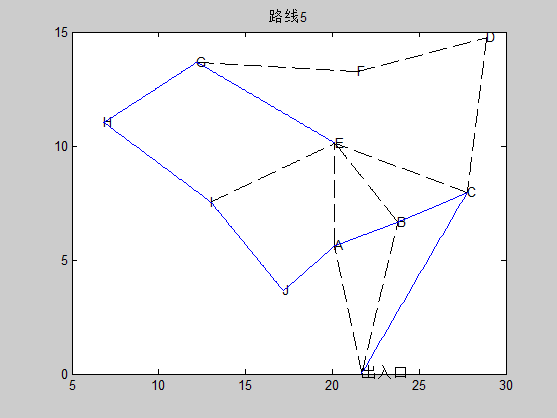
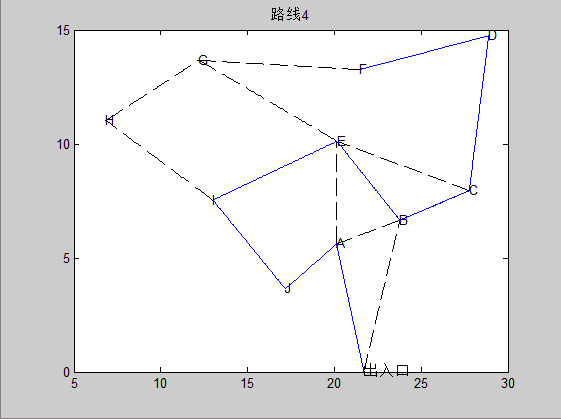
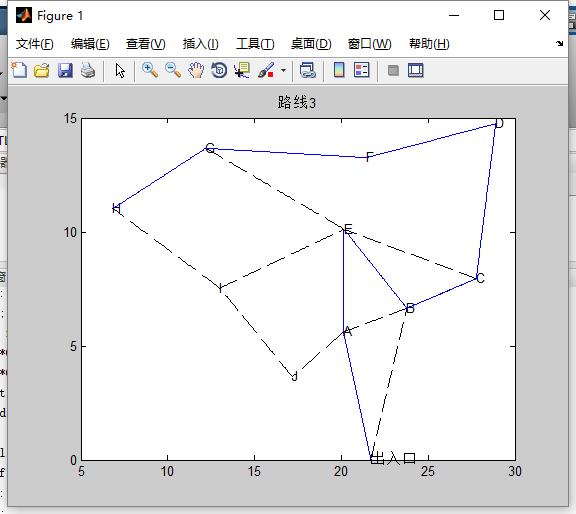
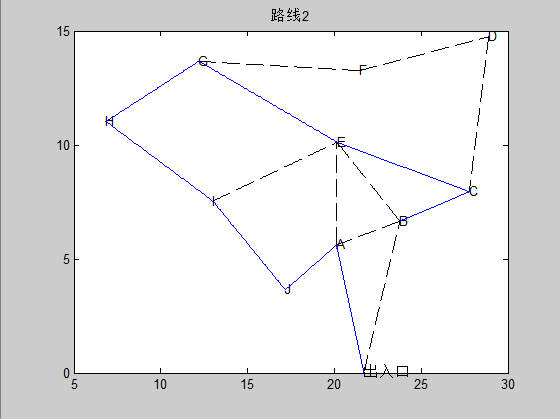
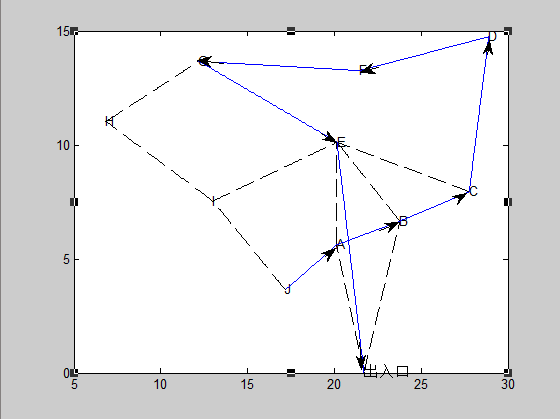
y=t\*C(E(m),2)+(1-t)\*C(E(m+1),2);

plot(x,y,'b')

hold on

end

**附录3 所有路线图**



**附录4 层次分析法代码**

%几何平均法（方根法）

n=length(B);

for i=1:n

d(i)=1;

for j=1:n

d(i)=d(i)\*B(i,j);

end

e(i)=d(i).^(1/n);

end

f=sum(e);

for i=1:n

W(i)=e(i)/f;

end

W

E=B\*W';

amax=0;

for i=1;n

amax=E(i)/(n\*W(i))

end

CI=(amax-n)/(n-1)

RI =[ 0 0 0.5200 0.8900 1.1200 1.2400 1.3600 1.4100 1.4600 1.4900 1.5200 1.5400 1.5600 1.5800 ];

CR=CI/RI(n)

%算数平均法，求和法

for i=1:n

c(i)=0;

for j=1:n

b=0;

for k=1:n

b=b+B(k,j);

end

c(i)=c(i)+B(i,j)/b;

end

W(i)=c(i)/n;

end

W

E=B\*W';

amax=0;

for i=1;n

amax=E(i)/(n\*W(i))

end

CI=(amax-n)/(n-1)

RI =[ 0 0 0.5200 0.8900 1.1200 1.2400 1.3600 1.4100 1.4600 1.4900 1.5200 1.5400 1.5600 1.5800 ];

CR=CI/RI(n)

**附录5 用matlab整理预定房间数据**

%先要excel数据导入

%A预定日期

%B到店的日期

%C离店的日期

%D每一单预定房间数

E=zeros(400,1);

a=735965;

>> l=length(B);

>> for m=1:l

d=C(m)-B(m)+1;

for i=1:d

E(B(m)-a+i)=E(B(m)-a+i)+D(m);

end

end

**附录6**

f(x) =

a1\*exp(-((x-b1)/c1)^2) + a2\*exp(-((x-b2)/c2)^2) +

a3\*exp(-((x-b3)/c3)^2) + a4\*exp(-((x-b4)/c4)^2) +

a5\*exp(-((x-b5)/c5)^2)

Coefficients (with 95% confidence bounds):

a1 = 194.8 (85.78, 303.8)

b1 = 335.8 (331.1, 340.4)

c1 = 36.11 (25.84, 46.38)

a2 = 68.57 (18.72, 118.4)

b2 = 276.6 (275.2, 277.9)

c2 = 2.299 (0.2933, 4.306)

a3 = 214.1 (119.1, 309.1)

b3 = 148.7 (140.1, 157.4)

c3 = 44.87 (31.76, 57.99)

a4 = 179.7 (138.6, 220.8)

b4 = 105.4 (103.9, 106.9)

c4 = 15.51 (12.08, 18.94)

a5 = 264.3 (253.4, 275.1)

b5 = 248.7 (240.3, 257.2)

c5 = 75.51 (30.7, 120.3)