

机票预售价格预测和数量决策的优化模型

摘要

本文针对航空公司机票预定价格的变动规律，以及在综合考虑经济利润和社会声誉情况下最优预售票数的决策进行了讨论。

在问题 1 中，本文建立了基于趋势比率法的季节预测模型和基于 SARIMA 的时间序列预测模型。首先根据价格走势初步认为原价格时间序列受长期趋势及季节变动的影响明显。之后由原数据建立直线趋势方程，得出各月的趋势值，并根据古典时间序列分解模型，分离出季节比率，之后调整季节比率得到季节指数，每月季节指数与各月的趋势值的乘积即为各月的机票价格的预测值。用这种方法可以得到 2010 年 4 月到 2011 年 3 月的机票预订价格预测值：1040、1106、1013、1138、1283、1032、1207、1080、935、1064、956、917，单位：元。在采用 SARIMA 方法建模时，用 EViews 软件首先通过一阶差分消除原序列的非平稳性，再进行一次季节性差分，削弱了季节因素的影响。然后考察平稳序列的自相关系数、偏相关系数，用 AIC 定阶准则建立了较为合适的 SARIMA 模型 $SARIMA(9,1,9)(1,12,1)^{12}$ ： $D_{12}DY_t = -6.676056 + u_t$

$$(1 + 0.810866L)(1 + 0.420978L^2)(1 - 0.404643L^9)(1 + 0.025949L^{12})u_t \\ = (1 - 1.004015L)(1 + 2.503009L^2)(1 + 5.261903L^9)(1 - 1.471241L^{12})v_t$$

并对其进行了白噪声检验。之后对模型进行样本内预测，通过 $MAPE$ 、 $Theil$ 、 BP 、 VP 、 CP 指标对预测精度进行了评估，结果显示较好。再次，对 2010 年 4 月到 2011 年 3 月机票预订价格进行预测，结果如下：886、961、830、972、1047、930、956、895、752、687、719、759，单位：元。最后针对两种方法的预测结果的差异进行分析，认为在预测时 SARIMA 模型较合适，并综合两种模型描述了机票价格的变动规律。

在问题 2 中，本文基于 Taguchi 模型，并做了一定的优化。以公司经济利益最大化和社会声誉尽量不受影响为原则公司。经济利益可以用机票收入扣除飞行费用和赔偿金，再加上订金收入后的利润来衡量，同时为了维持一定的社会声誉又要确保会被挤下飞机的乘客数量尽可能的小，因此建立了一个两目标的优化模型。而又考虑到预订票后乘客是否前来登机是随机的，所以文章运用概率的思想使其服从二项分布。公司的首先假定不优先考虑头等舱乘客的利益，分别对经济舱、头等舱用 MATLAB 进行了数值求解：当预订者不按时登机的概率 $p = 0.05$ 时，航空公司发出 318 张经济舱票、22 张头等舱票总收益大，单位费用的总收益 1.338；当预订者不按时登机的概率 $p = 0.1$ 时，航空公司发出 334 张经济舱票、23 张经济舱票总收益最大，单位费用收益也为 1.338。但又考虑到航空公司作为盈利组织，会充分保障头等舱乘客的利益，优先考虑这部分人的利益，以确保将来的客流量，所以本文从乘客转舱的角度，又分三种情况进行了更为深入探讨。

最后对以上三个模型进行了评价，以利于对其改进及推广。

【关键词】趋势比率法 SARIMA 模型 白噪声检验 Taguchi 模型 二项分布

一、问题重述

航空公司对机票一般采用预定策略。客户可以通过电话或互联网预定，这种预定具有很大的不确定性，客户很可能由于各种原因取消预定。航空公司为了争取更大的利润，一方面要争取客户，另一方面要降低客户取消预定遭受的损失。为此，航空公司采用一些措施。首先，要求客户提供信用卡号，预付一定数量的定金。如果客户在飞机起飞前48小时内取消预定，定金将如数退还，否则定金将被没收。其次，航空公司采用变动价格，根据市场需求情况调整机票价格，一般来说旺季机票价格比较高，淡季价格略低。

(1) 建立机票预定价格的数学模型，并对以下实例作分析。表中（见附录一）给出了某某航空公司某条航线2005年10月~2010年3月期间，每月经济舱机票平均价格（单位：元），利用所建模型说明价格变动的规律，并据此估计未来一年内的经济舱机票的参考价格。还可以收集更多的数据来佐证所建模型的价值（要求注明出处）。

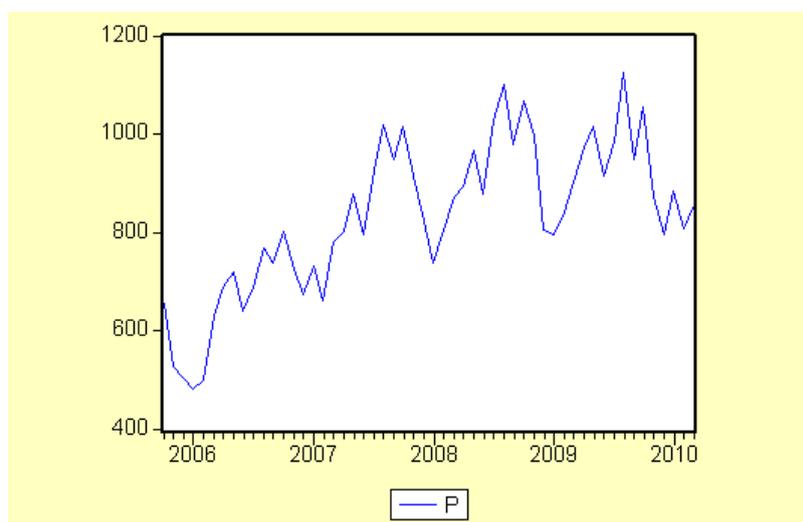
(2) 在旺季，航空公司往往可以预定出超过实际座位数的机票数，以减低客户取消预定航空公司的损失。当然这样做可能会带来新的风险，因为万一届时有超出座位数的客户出现，航空公司要通过升级机票档次或赔款来解决纠纷，为此航空公司还会承担信誉风险。某条航线就一中机型，有头等舱20座，经济舱300座，每天一班航班。试为该航线制定合理的预定策略，并论证理由。

二、问题的分析

2.1 问题一：

首先我们根据原始数据（机票价格）关于时间作出曲线图，如图1：

图1 机票价格关于时间的走势图



从该图我们初步判断，机票价格受到季节因子的影响，周期为一年，即12个月。每年的5、8、12月为旺季。同时，机票价格又受到长期趋势的影响，每年同期的价格呈上涨的趋势。

因此我们采用趋势比率法进行季节预测。建立趋势方程，并将长期趋势因子、季节变动因子分离，得到每一个月的季节比率。再算出一年后的趋势值，与季节比率的乘积即为之后一年的预测值。

另外我们建立 SARIMA 模型，特别适合于预测受季节因素影响的时间序列。

2.2 问题二:

由于航空公司订票业务的开展，若限制订票数额与飞机可容纳乘客数恰好相等，那么当乘客订了机票而未按时登机的情况发生时，飞机则会因为不满员而利润降低。若不限制订票数额，那么当前来登机的乘客超过飞机容量时，必然会引起部分乘客的抱怨，导致公司的声誉及经济造成损失。因此需要在二者辩证关系中，找到一个最佳订票数额点。

公司的经济利益可以用机票收入扣除飞行费用和赔偿金后的利润来衡量，声誉可以用按时前来登记但被挤下飞机的乘客限制在一定数量为标准^[1]。由于我们假定预订票的乘客是否前来登机是随机的，因此我们要讨论利润和声誉的平均期望值。该问题可以看做是一个最优化问题。

三. 模型假设

1. 在问题一中，本文假设不会出现突发事件导致机票价格或订票乘客数量有突然的改变。并且，在宏观、微观的政治经济环境上，将来一年不会发生巨大的变化。
2. 在旺季时，合理预定机票数量的限额都能达到，即所有预订票都能售出。
3. 假设每位预订票的乘客是否前来登机是互相独立的，并且是随机的。
4. 不优先考虑头等舱乘客的利益，即如果头等舱已经满员，多出的持头等舱票的乘客不能进入经济舱，只能等待下一班飞机并获得赔偿。

四、符号的约定

4.1 针对问题一:

t : 时间变量，单位：月；

T_t : 长期趋势因素；

S_t : 季节变动因素；

C_t : 周期变动因素；

I_t : 不规则变动因素；

Y_t : 原始时间序列，即机票价格序列；

\hat{Y}_t : 趋势值；

e : 残差；

a : 趋势方程的截距项。

b : 趋势方程的斜率

Δy_{t-i} : 时间序列的 i 阶差分，($i=0,1,2,\dots,m$)

d : 非季节差分的阶数

D : 季节差分的阶数

s : 季节周期

p : 非季节自相关的最大滞后阶数

q : 非季节偏自相关的最大滞后阶数

P : 季节自相关的最大滞后阶数

Q : 季节偏自相关的最大滞后阶数

4.2 针对问题二:

N : 飞机的容量

g : 机票的价格

f : 飞行费用

m : 预订票数量的限额

p : 每位乘客不按时前来登机的概率

b : 被挤掉后需付的赔偿金

S : 每次航班的利润

$p_j(m)$: 被挤掉的乘客人数超过 j 人的概率

$J(m)$: 单位费用获得的平均利润

\bar{S} : 平均利润

五、模型的建立与求解

5.1. 问题一模型的建立、求解与检验

5.1.1 模型一: 基于趋势比率法的季节预测模型

由古典时间序列分解模型我们知道, 时间序列可以分解成长期趋势 T_t 、季节变动 S_t 、周期变动 C_t 和不规则变动 I_t 四个因素。我们认为, 时间序列 Y 是这四个因素的函数, 即:

$$Y_t = f(T_t, S_t, C_t, I_t)$$

这个函数通常有两种形式:

1) 加法模型:
$$Y_t = T_t + S_t + C_t + I_t$$

2) 乘法模型:
$$Y_t = T_t * S_t * C_t * I_t$$

而选择加法模型的条件必须满足: 四个因素彼此独立。否则会产生重复计算问题。

而乘法模型不需要该条件。

因此针对本题，我们选择乘法模型计算。

1. 建立趋势方程，求各期趋势值。

令直线趋势方程为： $\hat{Y}_t = a + bt$ 。

根据标准最小二乘法原理：

$$e(a, b) = \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 = \sum_{t=1}^n (Y_t - a - bt)^2 = \text{最小值}$$

根据微分原理：

$$\begin{cases} \frac{\partial e}{\partial a} = 0 \\ \frac{\partial e}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \sum Y_t = na + b\sum t \\ \sum tY_t = a\sum t + b\sum t^2 \end{cases} \quad (1)$$

将数据代入试 (1) 得到：

$$\begin{cases} a = 636.9252 \\ b = 7.0882 \end{cases}$$

故，直线趋势方程为：

$$\hat{Y}_t = 636.9252 + 7.0882t \quad (2)$$

其中， t 的起始值为 1，该时间点位于 2005 年 10 月。

由于一般认为，趋势因素和循环因素不分开，而不规则变动因素是不可预测的，因此分离出循环因素和不规则变动因素对于时间序列的预测是没有多少价值的。故接下来我们直接分离季节因素。

2. 分离季节因素，求季节比率。

根据方程 (2)，即趋势方程，令 $t=1, 2, \dots, 54$ ，分别求出各年的趋势值。由季节比率的公式：季节比率=各期实际值/同期趋势值。求得各期的季节比率。见表 1：

表 1 机票价格直线趋势最小平方方法计算表：

| 年月 | 年次t | 价格Y | t*Y | t*t | 趋势值 | 季节比率 |
|----------|-----|------|-------|-----|------------|--------|
| 2005年10月 | 1 | 656 | 656 | 1 | 644.013468 | 1.0186 |
| 2005年11月 | 2 | 526 | 1052 | 4 | 651.101709 | 0.8079 |
| 2005年12月 | 3 | 502 | 1506 | 9 | 658.18995 | 0.7627 |
| 2006年1月 | 4 | 482 | 1928 | 16 | 665.278191 | 0.7245 |
| 2006年2月 | 5 | 498 | 2490 | 25 | 672.366432 | 0.7407 |
| 2006年3月 | 6 | 632 | 3792 | 36 | 679.454673 | 0.9302 |
| 2006年4月 | 7 | 688 | 4816 | 49 | 686.542914 | 1.0021 |
| 2006年5月 | 8 | 720 | 5760 | 64 | 693.631155 | 1.0380 |
| 2006年6月 | 9 | 640 | 5760 | 81 | 700.719396 | 0.9133 |
| 2006年7月 | 10 | 688 | 6880 | 100 | 707.807637 | 0.9720 |
| 2006年8月 | 11 | 768 | 8448 | 121 | 714.895878 | 1.0743 |
| 2006年9月 | 12 | 736 | 8832 | 144 | 721.984119 | 1.0194 |
| 2006年10月 | 13 | 802 | 10426 | 169 | 729.07236 | 1.1000 |
| 2006年11月 | 14 | 726 | 10164 | 196 | 736.160601 | 0.9862 |
| 2006年12月 | 15 | 672 | 10080 | 225 | 743.248842 | 0.9041 |
| 2007年1月 | 16 | 732 | 11712 | 256 | 750.337083 | 0.9756 |
| 2007年2月 | 17 | 662 | 11254 | 289 | 757.425324 | 0.8740 |
| 2007年3月 | 18 | 780 | 14040 | 324 | 764.513565 | 1.0203 |
| 2007年4月 | 19 | 802 | 15238 | 361 | 771.601806 | 1.0394 |
| 2007年5月 | 20 | 878 | 17560 | 400 | 778.690047 | 1.1275 |
| 2007年6月 | 21 | 794 | 16674 | 441 | 785.778288 | 1.0105 |
| 2007年7月 | 22 | 926 | 20372 | 484 | 792.8665 | 1.1679 |
| 2007年8月 | 23 | 1018 | 23414 | 529 | 799.9548 | 1.2726 |
| 2007年9月 | 24 | 948 | 22752 | 576 | 807.0430 | 1.1747 |
| 2007年10月 | 25 | 1016 | 25400 | 625 | 814.1313 | 1.2480 |

| | | | | | | |
|----------|----|------|-------|------|-----------|--------|
| 2007年11月 | 26 | 916 | 23816 | 676 | 821.2195 | 1.1154 |
| 2007年12月 | 27 | 824 | 22248 | 729 | 828.3077 | 0.9948 |
| 2008年1月 | 28 | 738 | 20664 | 784 | 835.3960 | 0.8834 |
| 2008年2月 | 29 | 806 | 23374 | 841 | 842.4842 | 0.9567 |
| 2008年3月 | 30 | 872 | 26160 | 900 | 849.5725 | 1.0264 |
| 2008年4月 | 31 | 894 | 27714 | 961 | 856.6607 | 1.0436 |
| 2008年5月 | 32 | 966 | 30912 | 1024 | 863.7489 | 1.1184 |
| 2008年6月 | 33 | 878 | 28974 | 1089 | 870.8372 | 1.0082 |
| 2008年7月 | 34 | 1028 | 34952 | 1156 | 877.9254 | 1.1709 |
| 2008年8月 | 35 | 1100 | 38500 | 1225 | 885.0137 | 1.2429 |
| 2008年9月 | 36 | 978 | 35208 | 1296 | 892.1019 | 1.0963 |
| 2008年10月 | 37 | 1068 | 39516 | 1369 | 899.1901 | 1.1877 |
| 2008年11月 | 38 | 996 | 37848 | 1444 | 906.2784 | 1.0990 |
| 2008年12月 | 39 | 804 | 31356 | 1521 | 913.3666 | 0.8803 |
| 2009年1月 | 40 | 794 | 31760 | 1600 | 920.4549 | 0.8626 |
| 2009年2月 | 41 | 832 | 34112 | 1681 | 927.5431 | 0.8970 |
| 2009年3月 | 42 | 902 | 37884 | 1764 | 934.6313 | 0.9651 |
| 2009年4月 | 43 | 972 | 41796 | 1849 | 941.7196 | 1.0322 |
| 2009年5月 | 44 | 1014 | 44616 | 1936 | 948.8078 | 1.0687 |
| 2009年6月 | 45 | 916 | 41220 | 2025 | 955.8961 | 0.9583 |
| 2009年7月 | 46 | 986 | 45356 | 2116 | 962.9843 | 1.0239 |
| 2009年8月 | 47 | 1124 | 52828 | 2209 | 970.0726 | 1.1587 |
| 2009年9月 | 48 | 948 | 45504 | 2304 | 977.1608 | 0.9702 |
| 2009年10月 | 49 | 1056 | 51744 | 2401 | 984.2490 | 1.0729 |
| 2009年11月 | 50 | 872 | 43600 | 2500 | 991.3373 | 0.8796 |
| 2009年12月 | 51 | 796 | 40596 | 2601 | 998.4255 | 0.7973 |
| 2010年1月 | 52 | 884 | 45968 | 2704 | 1005.5138 | 0.8792 |
| 2010年2月 | 53 | 808 | 42824 | 2809 | 1012.6020 | 0.7979 |
| 2010年3月 | 54 | 856 | 46224 | 2916 | 1019.6902 | 0.8395 |

3. 通过季节变动规律，计算季节指数

表 2 季节指数计算表

| 年月 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 合计 |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| 2005 | --- | --- | --- | --- | --- | --- | --- | --- | --- | 1.019 | 0.808 | 0.763 | --- |
| 2006 | 0.725 | 0.741 | 0.930 | 1.002 | 1.038 | 0.913 | 0.972 | 1.074 | 1.019 | 1.100 | 0.986 | 0.904 | --- |
| 2007 | 0.976 | 0.874 | 1.020 | 1.039 | 1.128 | 1.010 | 1.168 | 1.273 | 1.175 | 1.248 | 1.115 | 0.995 | --- |
| 2008 | 0.883 | 0.957 | 1.026 | 1.044 | 1.118 | 1.008 | 1.171 | 1.243 | 1.096 | 1.188 | 1.099 | 0.880 | --- |
| 2009 | 0.863 | 0.897 | 0.965 | 1.032 | 1.069 | 0.958 | 1.024 | 1.159 | 0.970 | 1.073 | 0.880 | 0.797 | --- |
| 2010 | 0.879 | 0.798 | 0.839 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 同月平均 | 0.865 | 0.853 | 0.956 | 1.029 | 1.088 | 0.973 | 1.084 | 1.187 | 1.065 | 1.125 | 0.978 | 0.868 | 12.071 |
| 季节指数 | 0.860 | 0.848 | 0.951 | 1.023 | 1.082 | 0.967 | 1.077 | 1.180 | 1.059 | 1.119 | 0.972 | 0.863 | 12.000 |
| 最近季节比率 | 0.879 | 0.798 | 0.839 | 1.032 | 1.069 | 0.958 | 1.024 | 1.159 | 0.970 | 1.073 | 0.880 | 0.797 | 11.478 |
| 季节指数 | 0.919 | 0.834 | 0.878 | 1.079 | 1.117 | 1.002 | 1.070 | 1.234 | 1.014 | 1.122 | 0.920 | 0.833 | 12.000 |
| 趋势分析比率 | 1.020 | 0.911 | 0.866 | 1.059 | 1.119 | 1.017 | 1.136 | 1.272 | 1.016 | 1.180 | 1.049 | 0.902 | 12.546 |
| 季节指数 | 0.975 | 0.871 | 0.828 | 1.013 | 1.070 | 0.973 | 1.086 | 1.216 | 0.972 | 1.128 | 1.004 | 0.863 | 12.000 |

我们取各年同月季节比率的平均得到同月平均季节比率，取最近一年内各月的季节比率作为最近季节比率，趋势分析比率=同月平均值+最近季节比率—最远季节比率。

同时由公式：

$$\frac{\text{季节比率}}{\text{季节指数}} = \frac{\text{季节比率合计}}{12}$$

且这 12 个月份季节指数之和为 12。以此得到季节指数。

4. 选取季节指数

从季节指数计算表（表 2）我们看到，各年同月季节指数从 2005 年到 2007 年呈增加趋势，而 2007 年到 2010 年呈减少趋势。因此要预测下一年的值，我们最好选取由趋势分析比率得到的季节指数，该指数适合季节变动大但趋势稳定的情况。

5. 根据乘法模型预测。

1) 首先我们算出下一年的趋势值：

令 $t = 55, 56, \dots, 66$ 。代入趋势方程（方程 2）中，得到：

$$Y_{55} = 1026.7785, Y_{56} = 1033.8667, Y_{57} = 1040.9550, Y_{58} = 1048.0432, \\ Y_{59} = 1055.1314, Y_{60} = 1062.2197, Y_{61} = 1069.3079, Y_{62} = 1076.3962, \\ Y_{63} = 1083.4844, Y_{64} = 1090.5727, Y_{65} = 1097.6609, Y_{66} = 1104.7491。$$

2) 计算各月趋势值与各月季节指数的乘积，如表 3:

该乘积即为下一年机票价格预测值。

表 3 下一年机票价格预测值

| 年月 | 趋势值 | 季节指数 | 预测值 |
|----------|-----------|--------|-----------|
| 2010年4月 | 1026.7785 | 1.0132 | 1040.3802 |
| 2010年5月 | 1033.8667 | 1.0702 | 1106.4073 |
| 2010年6月 | 1040.955 | 0.9732 | 1013.07 |
| 2010年7月 | 1048.0432 | 1.0862 | 1138.3431 |
| 2010年8月 | 1055.1314 | 1.2162 | 1283.2229 |
| 2010年9月 | 1062.2197 | 0.9717 | 1032.1223 |
| 2010年10月 | 1069.3079 | 1.1284 | 1206.6008 |
| 2010年11月 | 1076.3962 | 1.0037 | 1080.3906 |
| 2010年12月 | 1083.4844 | 0.8631 | 935.175 |
| 2011年1月 | 1090.5727 | 0.9753 | 1063.6571 |
| 2011年2月 | 1097.6609 | 0.8709 | 955.96941 |
| 2011年3月 | 1104.7491 | 0.8279 | 914.64256 |

5.1.2 模型二：基于 SARIMA 模型的时间序列预测

1. 数据的平稳性检验及平稳化

人们对时间序列分析是通过建立以因果关系为基础的结构模型进行的。这种分析背后有一个隐含的假设，即原始数据是平稳的。否则，通常的统计检验结果不可靠，甚至出现“虚假回归”问题，即两个毫无关系的非平稳时间序列很可能会得到一个“显著”的线性回归模型。因此，建立时间序列模型时首先需要检验时间序列数据的平稳性。^[1]

1) 平稳性的图示检验

由图 1(第 3 页), 机票价格没有围绕一个平均值上下波动, 而是有一个向上的趋势, 平且有明显的季节性, 周期为一年。

因此本题数据为包含趋势项的非平稳时间序列。

2) ADF 检验

ADF 检验从以下三个模型中选择一个模型进行检验:

$$\text{模型 1: } \Delta y_t = \delta y_{t-1} + \sum_{i=1}^m \beta_i \Delta y_{t-i} + \varepsilon_t$$

$$\text{模型 2: } \Delta y_t = \alpha + \delta y_{t-1} + \sum_{i=1}^m \beta_i \Delta y_{t-i} + \varepsilon_t$$

$$\text{模型 3: } \Delta y_t = \alpha + \beta t + \delta y_{t-1} + \sum_{i=1}^m \beta_i \Delta y_{t-i} + \varepsilon_t$$

其中模型 3 中其中 t 是时间变量, 反映了时间序列的趋势变化; 模型 2 包含了截距项, 而模型 1 不包含截距项和趋势项。

实际检验时原假设均为 $H_0: \delta = 0$, 即存在单位根。

用 EViews 对原始数据进行 ADF 检验, 从模型 3 开始, 若平稳 (拒绝原假设) 则停止, 若不平稳则按模型 2、模型 1 继续检验, 直到平稳为止。平稳的标志是 ADF 检验值小于软件输出的临界值。

过程如下:

1) 用原序列选择含截距项和趋势项的模型检验。输出结果如下:

| | | | |
|--------------------|-----------|--------------------|---------|
| ADF Test Statistic | -2.510668 | 1% Critical Value* | -4.1420 |
| | | 5% Critical Value | -3.4969 |
| | | 10% Critical Value | -3.1772 |

结果显示, ADF 检验值为-2.510668, 而显著水平取 1%、5%、10%的临界值分别是-4.1420、-3.4969、-3.1772。在任何一种显著水平下, ADF 检验值都大于临界值, 不能拒绝原假设, 该序列不平稳。

2) 接着用原序列选择仅含截距项的模型检验。输出结果如下:

| | | | |
|--------------------|-----------|--------------------|---------|
| ADF Test Statistic | -2.317841 | 1% Critical Value* | -3.5598 |
| | | 5% Critical Value | -2.9178 |
| | | 10% Critical Value | -2.5964 |

结果显示, ADF 检验值为-2.317841, 而显著水平取 1%、5%、10%的临界值分别是-3.5598、-2.9178、-2.5964。在任何一种显著水平下, ADF 检验值都大于临界值, 仍然不能拒绝原假设, 该序列不平稳。

3) 用原序列选择不含截距项和趋势项的模型检验。输出结果如下:

| | | | |
|--------------------|----------|--------------------|---------|
| ADF Test Statistic | 0.164623 | 1% Critical Value* | -2.6072 |
| | | 5% Critical Value | -1.9470 |
| | | 10% Critical Value | -1.6191 |

结果显示，ADF 检验值为 0.164623，而显著水平取 1%、5%、10%的临界值分别是 -32.6072、-1.9470、-1.6191。在任何一种显著水平下，ADF 检验值都大于临界值，依然不能拒绝原假设，该序列不平稳。

4) 在这种情况下我们将数据一阶差分。用一阶差分后的数据选择包含截距项和趋势项的模型检验。输出结果如下：

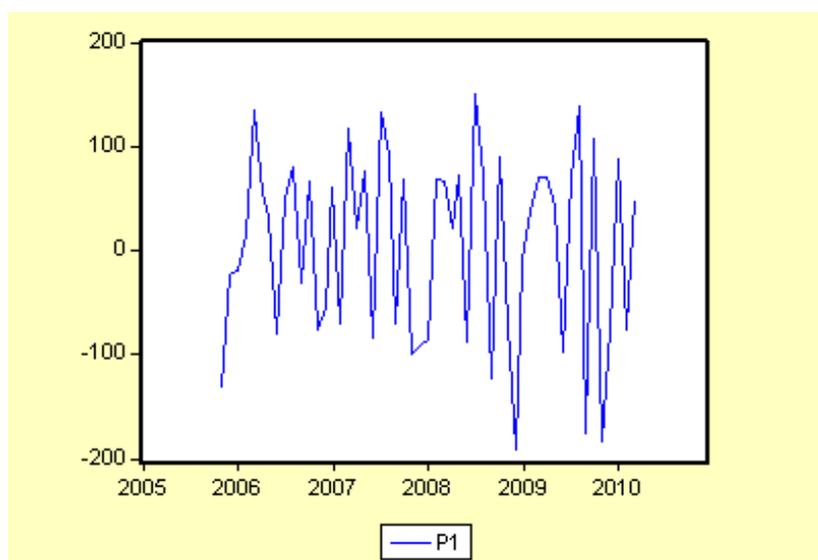
| | | | |
|--------------------|-----------|--------------------|---------|
| ADF Test Statistic | -5.859678 | 1% Critical Value* | -2.6081 |
| | | 5% Critical Value | -1.9471 |
| | | 10% Critical Value | -1.6191 |

结果显示，ADF 检验值为-25.859678，而显著水平取 1%、5%、10%的临界值分别是-2.6081、-21.9471、-1.6191。在任何一种显著水平下，ADF 检验值都小于临界值，因此拒绝原假设，该序列平稳。

2. 季节性差分处理

生成由原始序列 Y 经过一阶差分后形成的 Y1 序列，在 EVIEWS 命令窗口中输入：GENR Y1=d(Y,1)。接着输入：PLOT Y1，作出走势图，见图 2：

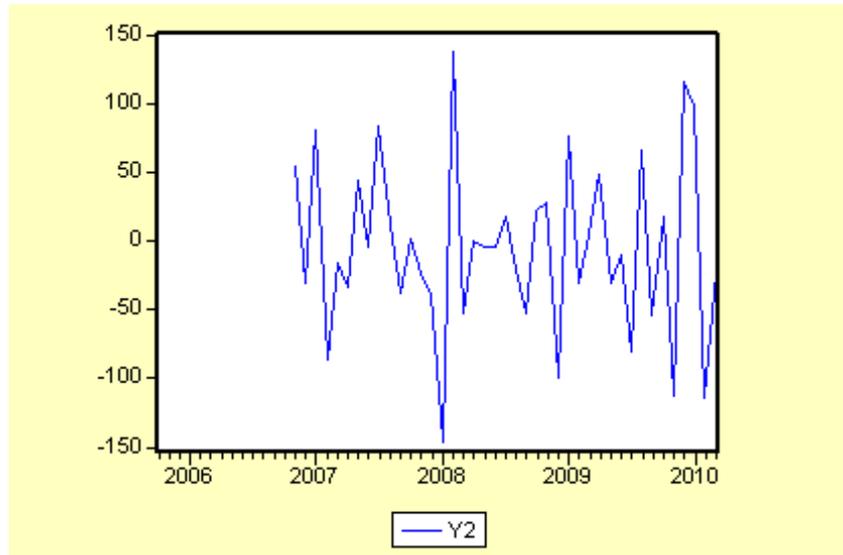
图 2 原数据一阶差分后关于时间走势图



我们看到，一阶差分消除了时间序列的趋势因子，但仍有明显的季节性，因此通过季节性差分来消除。

对一阶差分后的 Y 序列，再进行 12 阶差分，操作如下。在 EVIEWS 命令窗口中输入：GENR Y2=d(Y,1,12)，形成 Y2 序列。效果见图 3：

图 3 原数列一次差分和季节性差分后关于时间的走势图



很显然，季节因素被消除，数据围绕着一个平均值上下波动。对该序列进行 ADF 检验，认为序列是平稳的。输出结果如下：

| | | | |
|--------------------|-----------|--------------------|---------|
| ADF Test Statistic | -6.079702 | 1% Critical Value* | -3.6067 |
| | | 5% Critical Value | -2.9378 |
| | | 10% Critical Value | -2.6069 |

结果显示，ADF 检验值为-6.079702，而显著水平取 1%、5%、10%的临界值分别是-3.6067、-2.9378、-2.6069。在任何一种显著水平下，ADF 检验值都小于临界值，因此拒绝原假设，该序列平稳。

3. 模型的建立

在实现了时间序列的平稳化后，就要建立基于机票价格的 $SATIMA(p,d,q)(P,D,Q)^s$ 模型 这里涉及到 p 、 q 确定的问题。

由前面时间序列平稳化过程中进行的一阶差分、季节性差分，可以有 $d=1$ 、 $s=12$ 、 $D=1$ 。

依图 4，我们初步选定 (p, q) 的范围： p 可取 1、2、9、10、11， q 可取 1、2、6、9、11。在 EVIEWS 中对各种组合经过多次的试验，依据 AIC 最小的原则，最终选定 $(p, q)=(9,9)$ 。

图 4 相关系数及偏相关系数

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | |
|-----------------|---------------------|----|--------|--------|--------|
| | | 1 | -0.417 | -0.417 | 7.6811 |
| | | 2 | 0.014 | -0.194 | 7.6905 |
| | | 3 | -0.079 | -0.193 | 7.9830 |
| | | 4 | 0.063 | -0.073 | 8.1734 |
| | | 5 | -0.068 | -0.112 | 8.4001 |
| | | 6 | -0.125 | -0.276 | 9.1853 |
| | | 7 | 0.086 | -0.173 | 9.5721 |
| | | 8 | -0.051 | -0.222 | 9.7127 |
| | | 9 | 0.274 | 0.162 | 13.850 |
| | | 10 | -0.271 | -0.113 | 18.022 |
| | | 11 | 0.294 | 0.223 | 23.085 |
| | | 12 | -0.226 | -0.023 | 26.192 |
| | | 13 | 0.001 | -0.094 | 26.192 |
| | | 14 | 0.032 | 0.048 | 26.261 |
| | | 15 | 0.009 | 0.092 | 26.266 |
| | | 16 | -0.089 | -0.078 | 26.825 |
| | | 17 | -0.079 | -0.181 | 27.284 |
| | | 18 | 0.247 | -0.019 | 31.964 |
| | | 19 | -0.160 | -0.098 | 34.023 |
| | | 20 | 0.117 | -0.088 | 35.169 |

在 EViews 命令窗口中输入：LS Y2 C AR(1) AR(2) AR(9) MA(1) MA(2) MA(9) SAR(12) SMA(12)，得到输出结果：

图 5 EViews 估计结果

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| C | -6.676056 | 2.214003 | -3.015378 | 0.0118 |
| AR(1) | -0.810866 | 0.229413 | -3.534526 | 0.0047 |
| AR(2) | -0.420978 | 0.265965 | -1.582832 | 0.1418 |
| AR(9) | 0.404643 | 0.132761 | 3.047915 | 0.0111 |
| SAR(12) | -0.025949 | 0.411849 | -2.428077 | 0.0309 |
| MA(1) | -1.004015 | 0.415113 | -2.418654 | 0.0341 |
| MA(2) | 2.503009 | 0.640073 | 3.910504 | 0.0024 |
| MA(9) | 5.261903 | 1.241960 | 4.236772 | 0.0014 |
| SMA(12) | -1.471241 | 0.049960 | -2.001602 | 0.0662 |

其中 $R^2 = 0.934814$ 、 $\bar{R}^2 = 0.887407$ 、 $AIC = 9.377825$ 、 $SC = 9.825904$ 、 $Prob(F - statistic) = 0.000018 < 0.05$ 。说明模型拟合效果较好。

又 P 、 Q 为季节自回归、移动平均算子的最大滞后阶数，故 $P=1$ 、 $Q=1$ 。该模型是一个 $SARIMA(9,1,9)(1,12,1)^{12}$ 模型，即：

$$D_{12}DY_t = -6.676056 + u_t$$

$$(1 + 0.810866L)(1 + 0.420978L^2)(1 - 0.404643L^9)(1 + 0.025949L^{12})u_t$$

$$= (1 - 1.004015L)(1 + 2.503009L^2)(1 + 5.261903L^9)(1 - 1.471241L^{12})v_t$$

4. 模型的检验

模型的适应性检验：图 6 中，11 阶之后，各阶自相关与偏相关 Q 值对应的概率均大于 0.05，故认为该模型残差是白噪声，通过了该检验。

图 6 模型自相关与偏相关系数

| Autocorrelation | Partial Correlation | AC | PAC | Q-Stat | Prob | |
|-----------------|---------------------|----|--------|--------|--------|-------|
| | | 1 | 0.131 | 0.131 | 0.3998 | |
| | | 2 | 0.140 | 0.125 | 0.8782 | |
| | | 3 | -0.165 | -0.204 | 1.5801 | |
| | | 4 | -0.031 | -0.001 | 1.6057 | |
| | | 5 | -0.170 | -0.122 | 2.4571 | |
| | | 6 | -0.075 | -0.070 | 2.6348 | |
| | | 7 | -0.377 | -0.355 | 7.4380 | |
| | | 8 | 0.045 | 0.122 | 7.5138 | |
| | | 9 | -0.057 | -0.029 | 7.6444 | 0.006 |
| | | 10 | 0.102 | -0.062 | 8.0986 | 0.017 |
| | | 11 | 0.056 | 0.062 | 8.2502 | 0.041 |
| | | 12 | -0.028 | -0.178 | 8.2922 | 0.081 |
| | | 13 | -0.079 | -0.097 | 8.6803 | 0.123 |
| | | 14 | 0.046 | -0.037 | 8.8323 | 0.183 |
| | | 15 | -0.049 | -0.011 | 9.0449 | 0.249 |
| | | 16 | 0.022 | -0.049 | 9.0965 | 0.334 |
| | | 17 | -0.037 | -0.035 | 9.2926 | 0.411 |
| | | 18 | 0.025 | 0.037 | 9.4279 | 0.492 |

5. 预测

1) 样本内预测

由该模型预测 2008 年 8 月到 2010 年 3 月的机票价格。

表 4 2008 年 8 月~2010 年 3 月机票价格样本内预测值

| 年月 | 真实值 | 预测值 |
|----------|------|------|
| 2008年8月 | 1100 | 1083 |
| 2008年9月 | 978 | 1007 |
| 2008年10月 | 1068 | 1025 |
| 2008年11月 | 996 | 1012 |
| 2008年12月 | 804 | 838 |
| 2009年1月 | 794 | 774 |
| 2009年2月 | 832 | 845 |
| 2009年3月 | 902 | 885 |
| 2009年4月 | 972 | 922 |
| 2009年5月 | 1014 | 966 |
| 2009年6月 | 916 | 879 |
| 2009年7月 | 986 | 1007 |
| 2009年8月 | 1124 | 1103 |
| 2009年9月 | 948 | 958 |
| 2009年10月 | 1056 | 1012 |
| 2009年11月 | 872 | 987 |
| 2009年12月 | 796 | 785 |
| 2010年1月 | 884 | 742 |
| 2010年2月 | 808 | 785 |
| 2010年3月 | 856 | 827 |

表 4 样本内预测评估各检验值

| | |
|--------------------------------|----------|
| Root Mean Squared Error | 49.55532 |
| Mean Absolute Error | 36.93223 |
| Mean Absolute Percentage Error | 4.010556 |
| Theil Inequality Coefficient | 0.026511 |
| Bias Proportion | 0.068747 |
| Variance Proportion | 0.005915 |
| Covariance Proportion | 0.925338 |

表 4 显示, $MAPE$ 值为 4.010556, 它反应了预测的相对误差。一般认为 $MAPE$ 值低于 10 时, 模型的预测精度较高。 $Theil$ 不等系数为 0.026511, 它介于 0~1 之间, 数值越小, 预测值和真实值越接近。偏差率 BP 、方差率 VP 、协变量 CP 分别为 0.068747、0.005915、0.925338, 三者互相联系, 它们的取值通常在 0~1 之间, 且三者之和等于 1, 当模型较为理想时, 3 项误差主要集中在协变量 CP 上, 使其接近于 1, 而其他接近 0。

综上, 该模型预测的精度较高。我们可以利用它来预测下一年的机票价格。

2) 样本外预测

在 EVIEWS 中, 将时间范围扩大到 2011 年 3 月。利用 FORCAST 预测。结果如下:

表 5 将来一年的机票价格预测值

| 年月 | 预测值 |
|----------|------|
| 2010年4月 | 886 |
| 2010年5月 | 961 |
| 2010年6月 | 830 |
| 2010年7月 | 972 |
| 2010年8月 | 1047 |
| 2010年9月 | 930 |
| 2010年10月 | 956 |
| 2010年11月 | 895 |
| 2010年12月 | 752 |
| 2011年1月 | 687 |
| 2011年2月 | 719 |
| 2011年3月 | 759 |

5.1.3 模型一与模型二的分析对比

由表三（第9页）和表五（第15页）对比，明显看到模型一的预测值片高于模型二的预测值。

分析看出，模型一的预测值等于趋势值与季节指数的乘积，而建立直线趋势方程的做法默认了该时间序列的趋势呈直线上升的态势（因求得斜率大于0），而分析数据我们知道，09年后半年相比08年同期机票价格均有所下降，而并非直线上升，因此这里导致了一定的预测误差。另外，由季节比率的变化从05年~07年呈上升趋势，到了08年开始下降，因此季节比率的选择的科学性有待加强。并且，季节比率由真实值与趋势值相除得到，就再次引起了预测误差。

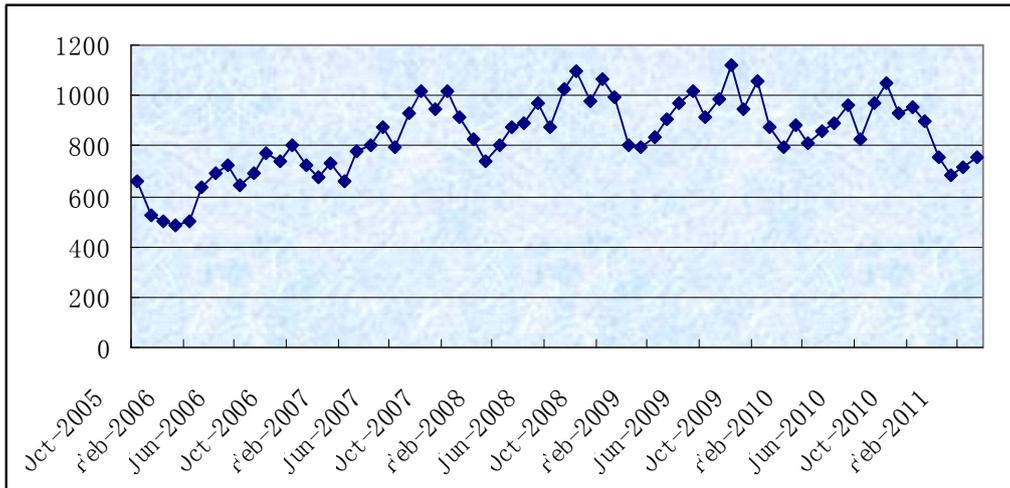
同时，对于模型二，在未来一年宏微观经济政治环境不会发生过大的变化的假设下，人们对交通的喜好、承受能力、评价延续前几年的规律。有上文分析，从2009年后半年开始该航空公司机票价格有所下降，于是这里我们认为，在2010年接下来的短期内，将继续保持这种趋势。利用SARIMA模型很好地反映了这一情况。从模型二的预测结果我们也可以看出，其值相对偏低。所以本文认为，模型二较好。

5.1.4 价格规律分析

SARIMA模型建立时进行了一阶差分、季节性差分，以消除时间序列的去趋势因素，使其围绕一个平均水平上下波动，以及消除了季节变动因素。由此看出，该航空公司机票价格包含一定的长期趋势，受季节影响，周期为1年。

研究原始数据以及对将来一年的预测值，自2005年3月至2009年7月，机票价格各年同期逐渐上升，而从2009年7月至2011年3月则有所下降。整体形成一个略微倒开口抛物线的趋势。以月为研究单位，每年1月价格较低，之后逐月上升，至5月达到第一个峰值，接着6月份下降一定幅度，7月份又开始增加，到8月达到第二个峰值，紧接着9月份价格降低少许，10月增加达到第三个峰值，之后呈下降趋势。在这三个峰值中，一般10月最高，5月最低。从实际来看，“五一”、“十一”黄金周，暑假是旅游旺季，人们出行的需求增加，机票供不应求，价格较高。

图7 各年月某航空公司机票价格走势



5.2 问题二模型的建立、求解与检验

5.2.1 模型三：优化的 Taguchi 模型

令飞机的容量为常数 N ，机票的价格为 g 。飞行费用 f 主要包括两大部分：一是飞行油耗；二是飞行中为乘客提供的服务费。一般来说，飞机费用与乘客数量关系很小，并且飞机费用可表示成机票价格 g 和飞机机身容量 N 的函数^[2]，即：

$$f = \alpha g N$$

预订票数量的限额为 $m (> n)$ ，每位乘客不按时前来登机的概率为 p ，并且假设每位乘客是否前来登机是互相独立的。如果有乘客预订了机票，但前来的登机的时候发现飞机已经满员，即自己被挤掉了，在这种情况下，航空公司可以让其机票升级或给予一定比例的赔偿金 b ，而赔偿金的金额与所预订的机票价格成正比，假定比例系数为 δ ，即 $b = \delta g$ 。公司为了降低顾客取消预订而带来的损失，该航空公司规定在预订机票的时候要交纳一定比例的定金，其与票价 g 成正比，比例系数为 β

模型建立与求解：

1) 为确定合理的预订票数量的限额，本文采用使航空公司的经济利益最大化的原则来衡量。每次航班的利润 S 应该由三部分组成^[8]：票房收益(订票取得的收益和取消订票对收益的影响)、运营成本和对被挤掉的乘客的赔偿。假定在这订票的 m 位乘客中有 k 位不按时前来登机时，航空公司将从飞行中得到的利润为：

$$S = \begin{cases} (m-k)g - f + \beta g(k - \eta_k), & m - k \leq N \\ Ng - f - (m - k - N)b + \beta g(k - \eta_k), & m - k > N \end{cases} \quad (3)$$

由假定，每位乘客不按时前来登机的概率为 p ，且由于每位乘客是否前来登记是相

互独立的，于是不按时前来登机的乘客数 k 服从二项分布^[9]，概率为：

$$P_k = C_m^k p^k q^{m-k} \quad (4)$$

当然，这种“未到”为相互独立的假定并不完全有效。事实上，部分乘客会成双或成群到达（或未到）。然而，让我们暂不考虑这种附带的困难，这样我们注意到：

$$\sum_{k=0}^m k p_k = m p$$

那么，航空公司由一次飞行获取的平均或期望利润为一个和式，它是所有可能的未到人对应情况下的利润乘以相应概率的和，即：

$$\bar{S} = \sum_{k=0}^m P_k \cdot [\text{由}(m-k)\text{个旅客带来的利润}] \quad (5)$$

$$= \sum_{k=0}^{m-N-1} P_k [(Ng - f) - (m - k - N)b + \beta g(k - \eta_k)] + \sum_{k=m-N}^m P_k [(m - k)g - f + \beta g(k - \eta_k)]$$

$$= \sum_{k=0}^{m-N-1} P_k [(N - m + k)g - (m - k - N)b] + (mg - f) \sum_{k=0}^m P_k - g \sum_{k=0}^m k P_k + \beta g \sum_{k=0}^m (k - \eta_k) P_k$$

而 $\sum_{k=0}^m P_k = 1$ 且 $\bar{k} = \sum_{k=0}^m k p_k = mp$ 是不按时到达的乘客数的数学期望值，用 \bar{k} 来表示，那么，

$$\begin{aligned} \bar{S} &= mg - f - \bar{k}g + \beta g \sum_{k=0}^m (k - \eta_k) P_k - (b + g) \sum_{k=0}^{m-N-1} P_k (m - N - k) \\ &= (m - \bar{k})g - f + \beta g \sum_{k=0}^m (k - \eta_k) P_k - (b + g) \sum_{k=0}^{m-N-1} P_k (m - N - k) \\ &= qmg - f + \beta g \sum_{k=0}^m (k - \eta_k) P_k - (b + g) \sum_{k=0}^{m-N-1} P_k (m - N - k) \end{aligned}$$

我们需要通过该式求得 \bar{S} 的最大值。

2) 航空公司为了未来的长远发展，必须考虑公司的社会声誉和经济利益，所以应该要求被挤掉的乘客人数不能太多。但是被挤掉者的数量是随机的，可以用被挤掉的乘客数超过若干人的概率作为度量指标。记被挤掉的乘客人数超过 j 人的概率为 $p_j(m)$ ，因为被挤掉的乘客超过 j 人，等价于 m 位预订票的乘客中不按时前来登机的不超过 $m - N - j - 1$ 人，所以就有：

$$p_j(m) = \sum_{k=0}^{m-N-j-1} P_k$$

对于给定的 N, j , 显然当 $m = N + j$ 时被挤掉的乘客不会超过 j 人, 即其概率为 0。而当 m 变大时 $p_j(m)$ 单调增加。

综上所述, 航空公司必须保证 $p_j(m)$ 的值不超过某一要求的定值。因此该问题可以描述成以利润最大化为一个约束条件, 平均利润 \bar{S} 为目标函数的最优化问题。

3) 我们可以通过飞行费用 $f = \alpha g N$ 化解平均利润。在平均利润的两边都同时除以飞行费用 f 为新的目标函数 $J(m)$, 其含义为单位费用获得的平均利润。所以可得:

$$J(m) = \bar{S} / f = \frac{1}{\alpha N} [qm + \beta \sum_{k=0}^m (k - \eta_k) P_k - (1 + b/g) \sum_{k=0}^{m-N-1} P_k(m - N - k)] - 1$$

其中 b/g 是对被挤掉的人的赔偿金占机票价格中的比例。所以上述问题可以化为:

$$\max J(m) = \frac{1}{\alpha N} [qm + \beta \sum_{k=0}^m (k - \eta_k) P_k - (1 + b/g) \sum_{k=0}^{m-N-1} P_k(m - N - k)] - 1$$

$$p_j(m) = \sum_{k=0}^{m-N-j-1} p_k \leq \omega$$

$$0 < \omega < 1$$

该模型无法求得解析解^[10], 所以本文设定几组数据, 用 matlab 软件作数值计算。

4) 分经济舱和头等舱求解

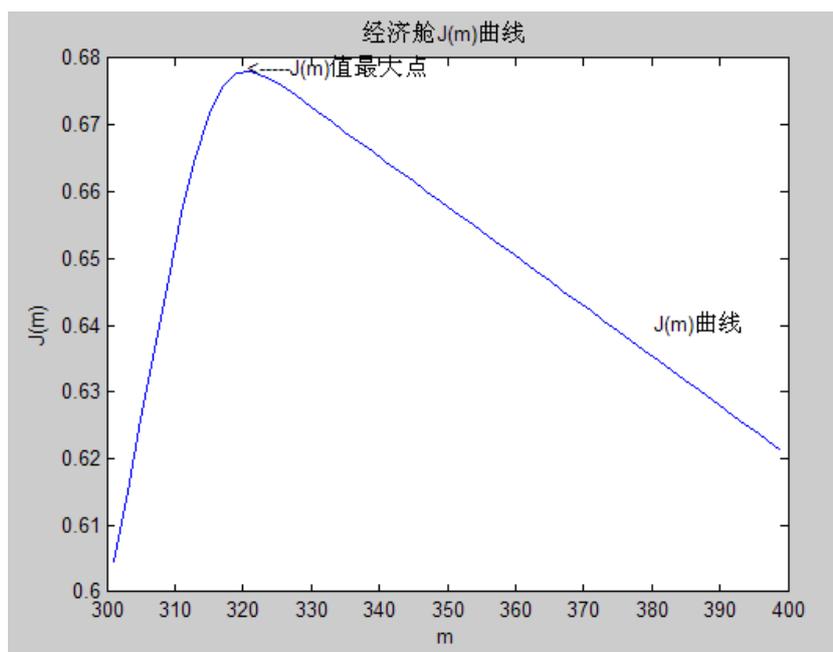
根据假设不优先考虑头等舱乘客的利益, 即如果头等舱已经满员, 多出的持头等舱票的乘客不能进入经济舱, 只能等待下一班飞机并获得赔偿。根据这个假设, 我们便可以将经济舱和头等舱分开来考虑。

a. 经济舱:

对于经济舱, 我们采用参考文献中的相关数据^[1]: $\alpha = 0.6$, $p = 0.05$, $b/g = 0.2$,

而定金比例 $\beta = 0.2$ 。考虑到在不按时前来登机的乘客中, 不取消预订的人数很小, 所以在此假定该人数占不按时来登机的人数的比例为 0.05, 而经济舱的座位数 $N = 300$, 使用 matlab 软件编程 (见附录二) 求解得到如下的 $J(m)$ 与 m 的关系图:

图 8 $J(m)$ 与 m 关系图

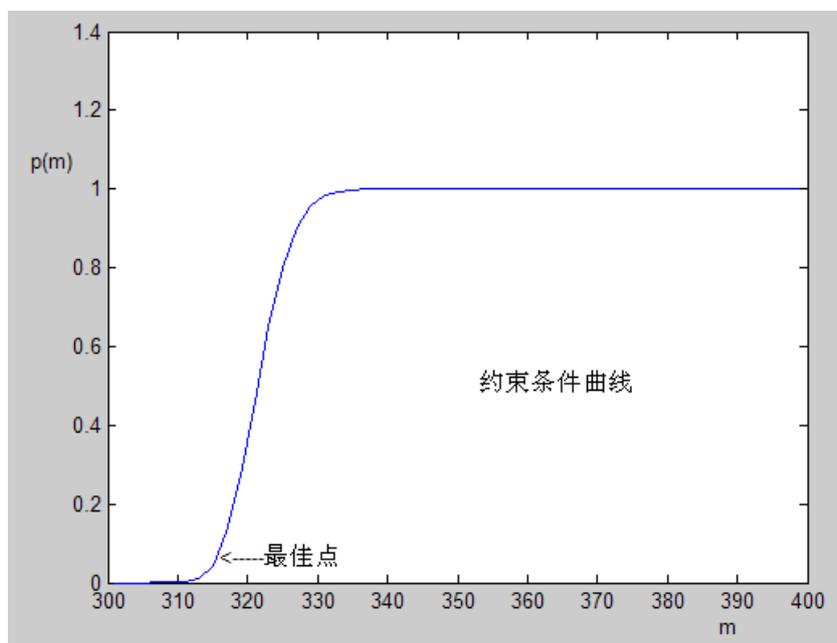


从上图中可以发现，在 $m = 320$ 时， $J(m)$ 取得最大值，但这还少考虑了其中的约束条件：

$$p_j(m) = \sum_{k=0}^{m-N-j-1} p_k \leq \omega$$

同样使用 matlab 软件编程（见附录二）求出在不同的 m 值下， $p_j(m)$ 的变化情况：

图 9



由图 9 可以看出，在 $m = 320$ 时，被挤掉的人数是 5 人的概率为 0.3627，在实际生活中这个值偏大，所以取 $m = 320$ 并不是很合适。使用 matlab 软件编程，给定相关参数，

对不同的 m 值，分别在不按时前来登机的概率 $p=0.05$ 和 $p=0.1$ 下，计算 $J(m)$ ， $p_5(m)$ 和 $p_{10}(m)$ ，从而得到一组表格（见附录三，表 6）

结论：综合考虑航空公司的经济效益和社会声誉，本文给定 $p_5(m) < 0.2$ ， $p_{10}(m) < 0.05$ ，由表 6 知，对于经济舱座位数 $N=300$ ，若估计 $p=0.05$ 时，取 $m=318$ ，使得在约束条件下，单位费用所产生的利益 $J_{\text{经}}(m) = 0.6769$ 为最大；若估计 $p=0.1$ 时，取 $m=334$ ，此时单位费用所产生的利益 $J_{\text{经}}(m) = 0.6899$ 为最大。

b. 头等舱：

求解过程和经济舱一样，但这里取座位数 $N=20$ ，同样使用 matlab 软件编程在不同的 m 值下，得到表 7（见附录四）。

结论：为综合考虑公司的利益，我们使得 $p_1(m) < 0.4$ ， $p_2(m) < 0.2$ ，再根据表中数据得：由头等舱座位数 $N=20$ ，若估计 $p=0.05$ 时，取 $m=22$ ，使得在约束条件下，单位费用所产生的利益 $J_{\text{头}}(m) = 0.6611$ 为最大。若估计 $p=0.1$ 时，取 $m=23$ ，使得在约束条件下，单位费用所产生的利益 $J_{\text{头}}(m) = 0.6699$ 最大。

综上，保证被挤掉的乘客人数较小，当预订者不按时前来登机的概率 $p=0.05$ 时，航空公司在发出 318 张经济舱票，22 张头等舱票时总收益最大，单位费用总的收益为：

$$J_{\text{总}}(m) = J_{\text{经}}(m) + J_{\text{头}}(m) = 0.6769 + 0.6611 = 1.338;$$

当预订者不按时前来登机的概率 $p=0.1$ 时，航空公司则在发出 334 张经济舱票，23 张头等舱票时总收益最大，单位费用总的收益为：

$$J_{\text{总}}(m) = J_{\text{经}}(m) + J_{\text{头}}(m) = 0.6899 + 0.6699 = 1.338$$

5.2.2 模型的深入

在模型三中假设该航空公司不充分保证头等舱的乘客的利益，即如果头等舱已经满员，多出的持头等舱预定票的乘客不能进入经济舱，只能等待下一班飞机并获得赔偿。但实际生活中，航空公司不可能愿意这么做的，因为头等舱的乘客属于高消费者，对于航空公司来说，从他们身上取得的利益远比经济舱的低消费者多，所以航空公司很不愿增加这部分人的抱怨率。为了更加的贴近实际，我们对于转舱作如下规定^[12]：

1. 如果头等舱已满，多出来的持该舱票的乘客优先享受次一级的待遇并得到赔偿，即如果一个持头等舱的乘客发现头等舱已满，那么他会被安排到经济舱而不会被挤下飞机，当然要满足这一点，很可能要挤下去一个经济舱的乘客。

2. 而当经济舱已满，而头等舱没有满员时，持经济舱票而被挤下飞机的乘客可以有机票升级的机会，即可以去乘坐头等舱作为补偿；但如果头等舱已经满员，则只能等下一班飞机并获得赔偿。

现在假设头等舱、经济舱的座位个数分别为 N_1, N_2 ，并且 $N_1 < N_2$ ，票价分别为 g_1, g_2 ，航空公司对两者预订票数量的限额分别为 m_1, m_2 ，两者不按时到达的人数分别为 k_1, k_2 ，为了叙述方便，把两者的飞行费用分开计算，分别为 f_1, f_2 。

这里对模型三作进一步优化^[13]，分成三种情况考虑：

1. 头等舱刚好满员。对于这种情形，也就不存在转舱的可能，所以与前面模型三的求解一致。

2. 头等舱已经满员。假设持头等舱票而被挤出头等舱的人数为 a ，由航空公司转舱的规定可知，这部分人会被安排到经济舱。而由假设知在旺季的时机票总是会被预定完的。并且为简单考虑，我们假设持头等舱票的乘客因头等舱满员而转到经济舱的，航空公司对其的补偿为两者票价的价差，即 $g_1 - g_2$ ，所以对于模型三的利润函数作如下修改：

$$S_2 = \begin{cases} (a + N_1)g_1 - f_1 - (m_1 - k_1 - N_1)(g_1 - g_2) + \beta g_1(k_1 - \eta_{k_1}) \\ + (m_2 - k_2)g_2 - f_2 + \beta g_2(k_2 - \eta_{k_2}), \\ (a + N_1)g_1 - f_1 - (m_1 - k_1 - N_1)(g_1 - g_2) + \beta g_1(k_1 - \eta_{k_1}) \\ + (N_2 - a)g_2 - f_2 - (m_2 - k_2 - N_2 + a)b + \beta g_2(k_2 - \eta_{k_2}), \end{cases}$$

$$= \begin{cases} N_1g_1 - f_1 + ag_2 + \beta g_1(k_1 - \eta_{k_1}) \\ + (m_2 - k_2)g_2 - f_2 + \beta g_2(k_2 - \eta_{k_2}), & m_2 - k_2 \leq N_2 - a \\ N_1g_1 - f_1 + ag_2 + \beta g_1(k_1 - \eta_{k_1}) \\ + (N_2 - a)g_2 - f_2 - (m_2 - k_2 - N_2 + a)b + \beta g_2(k_2 - \eta_{k_2}), & m_2 - k_2 > N_2 - a \end{cases}$$

接着与模型三相同的方法化简并求解。

3. 头等舱还有空位存在。假设头等舱存在的空位数为 d ，则根据规定知航空公司会把持经济舱票而被挤出的乘客升级到头等舱，作为补偿。所以在这种情形下，模型三的利润函数又可以修改为：

$$S_2 = \begin{cases} (m_1 - k_1)g_1 - f_1 + \beta g_1(k_1 - \eta_{k_1}) \\ + (m_2 - k_2)g_2 - f_2 + \beta g_2(k_2 - \eta_{k_2}), & m_2 - k_2 \leq N_2 + N_1 - m_1 - k_1 \\ (m_1 - k_1)g_1 - f_1 + \beta g_1(k_1 - \eta_{k_1}) \\ + (m_2 - k_2)g_2 - f_2 + \beta g_2(k_2 - \eta_{k_2}) + (m_2 - k_2 - N_2 - N_1 + m_1 + k_1)b, & m_2 - k_2 > N_2 + N_1 - m_1 - k_1 \end{cases}$$

当只有两个机舱时，所有可能发生的情况都包含在上面三种模型中，经过与模型三相同的化解后，并运用相应的软件求解，能够得出比模型三更加符合实际情形的解。

六、模型的评价：

6.1 优点:

1. 对于模型一，将趋势因子与季节因子分离，利于分别讨论各自对飞机票价的影响。可根据季节变动的具体情况，选择合适的季节比率，以此计算季节指数，预测结果更合理。计算简单，易操作。

2. 对于模型二，SARIMA 模型如今已十分的成熟，它的检验准则和参数估计方法可以得到对原序列较为精确的拟合。该模型的适用范围较广，许多种波动规律的序列均可。并且不少软件中，该模型都可以方便的实现。

3. 对于模型三，在保证被挤掉乘客数量较低的情况下，使得单位费用获得的平均利润最大而建立了优化的 Taguchi 模型，对模型进行一定的优化后，采用 matlab 软件求解，并绘出相应的图，使得分析更加清晰。同时，对模型进行了深化，从航空公司利益出发分三个情况考虑了乘客转舱的问题，从而更符合实际。

4. 对于模型三，建模时假设不按时前来登机的乘客数服从二项分布，考虑到了概率问题；在求解时又对不同的概率都作了相应的比较分析，使得结论更加有说服力。

6.2 缺点与改进:

1. 在用季节预测模型时，时间序列期数不完全可能导致对模型选择的误判，其中也包括对趋势方程形式的误判。可以通过对序列的环比发展速度、一次增长量的环比发展速度进行分析，判别时间序列的形式。

2. SARIMA 模型的定阶是难点，需要通过增加经验来获得更合理的模型。SARIMA 模型对于短期的预测效果较好，而期限过长，外推的精度逐渐减弱。但本文是对将来一年的预测，因此还是认为模型的预测效果较好。

3. 针对模型三，模型深化中虽然对乘客转舱问题作了全面的分析，但如果飞机机舱数目增多时，这个模型就会变的相对比较复杂，所以要进行推广时会碰到一些问题。

七、参考文献

- 【1】 姜启源、谢金星、叶俊，数学模型。北京：高等教育出版社，2003年
- 【2】 刘莺, 栗莉萍, 浅析航空公司的预订票策略, 科技创新导报, 2008. 36
- 【3】 徐国祥, 统计预测与决策, 第三版。上海：上海财经大学出版社，2008年
- 【4】 赵卫亚、彭寿康、朱晋，计量经济学。北京：机械工业出版社，2008年
- 【5】 李国柱、刘德智、赖志花、李从欣，计量经济学实验教程。北京：中国经济出版社，2010年
- 【6】 夏天、程细玉，SARIMA模型的建模及其信贷预测分析，华侨大学学报，第27卷第3期，2006年7月：229页到232页
- 【7】 吴玉清、孟昭时、郭晓江，机票预售价格和策略的数学模型，<http://www.docin.com/p-55324661.html>，2010年8月9日。
- 【8】 彭斯俊, 万丽军, 唐涛等, 基于演化计算的机票超售建模[J], 武汉理工大学学报(信息与管理工程版), 430070
- 【9】 商桂娥, 徐月芳, 基于二项式分布的多航班超售模型研究, 南京航空航天大学民航学院, 210016
- 【10】 张兴永, 周圣武, 朱开永, 机票预订问题的多阶段决策模型, 中国矿业大学数学系, 221008
- 【11】 徐瑞, 黄兆东, 阎凤玉, MATLAB科学计算与工程分析。科学出版社, 2007年
- 【12】 朱道元, 数学建模精品案例[M]。南京:东南大学出版社, 1999年
- 【13】 衡红军, 航班座位超售量的确定[J], 计算机工程, 2005, 31(7):166-167

八、附录

附录一：数据来源

某航空公司某条航线 2005 年 10 月~2010 年 3 月经济舱月平均价格(单位:元)

| 时间 | 价格 | 时间 | 价格 | 时间 | 价格 |
|----------|-----|----------|------|----------|------|
| 2005. 10 | 656 | 2007. 04 | 802 | 2008. 10 | 1068 |
| 2005. 11 | 526 | 2007. 05 | 878 | 2008. 11 | 996 |
| 2005. 12 | 502 | 2007. 06 | 794 | 2008. 12 | 804 |
| 2006. 01 | 482 | 2007. 07 | 926 | 2009. 01 | 794 |
| 2006. 02 | 498 | 2007. 08 | 1018 | 2009. 02 | 832 |
| 2006. 03 | 632 | 2007. 09 | 948 | 2009. 03 | 902 |
| 2006. 04 | 688 | 2007. 10 | 1016 | 2009. 04 | 972 |
| 2006. 05 | 720 | 2007. 11 | 916 | 2009. 05 | 1014 |
| 2006. 06 | 640 | 2007. 12 | 824 | 2009. 06 | 916 |
| 2006. 07 | 688 | 2008. 01 | 738 | 2009. 07 | 986 |
| 2006. 08 | 768 | 2008. 02 | 806 | 2009. 08 | 1124 |
| 2006. 09 | 736 | 2008. 03 | 872 | 2009. 09 | 948 |
| 2006. 10 | 802 | 2008. 04 | 894 | 2009. 10 | 1056 |
| 2006. 11 | 726 | 2008. 05 | 966 | 2009. 11 | 872 |
| 2006. 12 | 672 | 2008. 06 | 878 | 2009. 12 | 796 |
| 2007. 01 | 732 | 2008. 07 | 1028 | 2010. 01 | 884 |
| 2007. 02 | 662 | 2008. 08 | 1100 | 2010. 02 | 808 |
| 2007. 03 | 780 | 2008. 09 | 978 | 2010. 03 | 856 |

附录二: Matlab 程序

程序一:

```
num=1;
title('经济舱 J(m)曲线');xlabel('m');ylabel('J(m)');
p=0.05; %每位乘客不按时前来登机的概率
a=0.6; %飞机费用与机票价格的比例系数
n=300; %飞机上的座位数
b=0.2; %订票时所需要交的定金
c=0.2; %对被挤掉的乘客负的赔偿金与机票的比例
for m=300:2:401;
s(num)=1/(a*n)*((1-p)+0.05*b*p)*m-1;
for k=0:m-n-1
s(num)=s(num)-1/(a*n)*((1-p)+0.05*b*p)*(1+c)*(m-n-k)*binopdf(k,m,p);
%每单位费用所取得的收益
end
num=num+1;
end
m=300:2:401;
plot(m,s,'-')
```

程序二:

```
num=1;
title('约束曲线');xlabel('m');ylabel('p(m)');
p=0.05; %每位乘客不按时前来登机的概率
j=5; %被挤掉人数
for m=300:2:401; p(num)=0;
for k=0:m-n-j-1
p(num)=p(num)+binopdf(k,m,p); %计算有 j 人被挤掉的概率
end
num=num+1;
end
m=300:2:401;
plot(m,p,'-')
```

附录三(表 6):

| m | p=0.05 | | | | p=0.1 | | | |
|-----|---------|---------|--------|--------|---------|---------|--------|--------|
| | J | | P5 | P10 | J | | P5 | P10 |
| | b/g=0.2 | b/g=0.4 | | | b/g=0.2 | b/g=0.4 | | |
| 300 | 0.5992 | 0.5992 | 0.0000 | 0.0000 | 0.5317 | 0.5317 | 0.0000 | 0.0000 |
| 302 | 0.6098 | 0.6098 | 0.0000 | 0.0000 | 0.5419 | 0.5419 | 0.0000 | 0.0000 |
| 304 | 0.6205 | 0.6205 | 0.0000 | 0.0000 | 0.5521 | 0.5521 | 0.0000 | 0.0000 |
| 306 | 0.6311 | 0.6311 | 0.0000 | 0.0000 | 0.5623 | 0.5623 | 0.0000 | 0.0000 |
| 308 | 0.6417 | 0.6417 | 0.0000 | 0.0000 | 0.5725 | 0.5725 | 0.0000 | 0.0000 |
| 310 | 0.6519 | 0.6518 | 0.0005 | 0.0000 | 0.5827 | 0.5827 | 0.0000 | 0.0000 |
| 312 | 0.6611 | 0.6608 | 0.0044 | 0.0000 | 0.5929 | 0.5929 | 0.0000 | 0.0000 |
| 314 | 0.6686 | 0.6678 | 0.0232 | 0.0001 | 0.6031 | 0.6031 | 0.0000 | 0.0000 |
| 316 | 0.6740 | 0.6722 | 0.0791 | 0.0013 | 0.6134 | 0.6134 | 0.0000 | 0.0000 |
| 318 | 0.6769 | 0.6739 | 0.1931 | 0.0092 | 0.6235 | 0.6235 | 0.0000 | 0.0000 |
| 320 | 0.6780 | 0.6734 | 0.3627 | 0.0396 | 0.6337 | 0.6337 | 0.0002 | 0.0000 |
| 322 | 0.6778 | 0.6714 | 0.5558 | 0.1159 | 0.6437 | 0.6437 | 0.0008 | 0.0000 |
| 324 | 0.6769 | 0.6685 | 0.7295 | 0.2521 | 0.6535 | 0.6534 | 0.0030 | 0.0001 |
| 326 | 0.6756 | 0.6653 | 0.8565 | 0.4340 | 0.6628 | 0.6626 | 0.0093 | 0.0003 |
| 328 | 0.6742 | 0.6618 | 0.9335 | 0.6231 | 0.6714 | 0.6709 | 0.0243 | 0.0012 |
| 330 | 0.6727 | 0.6583 | 0.9730 | 0.7807 | 0.6789 | 0.6779 | 0.0547 | 0.0042 |
| 332 | 0.6712 | 0.6548 | 0.9903 | 0.8885 | 0.6851 | 0.6835 | 0.1074 | 0.0123 |
| 334 | 0.6697 | 0.6513 | 0.9969 | 0.9504 | 0.6899 | 0.6874 | 0.1869 | 0.0307 |
| 336 | 0.6683 | 0.6478 | 0.9991 | 0.9805 | 0.6933 | 0.6896 | 0.2922 | 0.0660 |
| 338 | 0.6668 | 0.6443 | 0.9998 | 0.9932 | 0.6954 | 0.6904 | 0.4158 | 0.1248 |
| 340 | 0.6653 | 0.6407 | 1.0000 | 0.9979 | 0.6966 | 0.6900 | 0.5455 | 0.2103 |
| 342 | 0.6638 | 0.6372 | 1.0000 | 0.9994 | 0.6969 | 0.6887 | 0.6680 | 0.3199 |

附录四（表 7）：

| m | p=0.05 | | | | p=0.1 | | | |
|----|---------|---------|--------|--------|---------|---------|--------|--------|
| | J | | P1 | P2 | J | | P1 | P2 |
| | b/g=0.2 | b/g=0.4 | | | b/g=0.2 | b/g=0.4 | | |
| 20 | 0.5992 | 0.5992 | 0.0000 | 0.0000 | 0.5317 | 0.5317 | 0.0000 | 0.0000 |
| 21 | 0.6465 | 0.6410 | 0.0000 | 0.0000 | 0.5982 | 0.5965 | 0.0000 | 0.0000 |
| 22 | 0.6611 | 0.6447 | 0.3235 | 0.0000 | 0.6446 | 0.6379 | 0.0985 | 0.0000 |
| 23 | 0.6585 | 0.6284 | 0.6794 | 0.3074 | 0.6699 | 0.6547 | 0.3151 | 0.0886 |
| 24 | 0.6497 | 0.6048 | 0.8841 | 0.6608 | 0.6797 | 0.6534 | 0.5643 | 0.2925 |
| 25 | 0.6390 | 0.5790 | 0.9659 | 0.8729 | 0.6806 | 0.6416 | 0.7636 | 0.5371 |
| 26 | 0.6279 | 0.5528 | 0.9915 | 0.9613 | 0.6773 | 0.6250 | 0.8882 | 0.7409 |
| 27 | 0.6168 | 0.5264 | 0.9981 | 0.9900 | 0.6721 | 0.6062 | 0.9529 | 0.8734 |
| 28 | 0.6056 | 0.5000 | 0.9996 | 0.9977 | 0.6663 | 0.5867 | 0.9821 | 0.9450 |
| 29 | 0.5944 | 0.4736 | 0.9999 | 0.9995 | 0.6603 | 0.5669 | 0.9938 | 0.9784 |
| 30 | 0.5832 | 0.4473 | 1.0000 | 0.9999 | 0.6542 | 0.5470 | 0.9980 | 0.9922 |
| 31 | 0.5720 | 0.4209 | 1.0000 | 1.0000 | 0.6481 | 0.5271 | 0.9994 | 0.9974 |
| 32 | 0.5608 | 0.3945 | 1.0000 | 1.0000 | 0.6420 | 0.5072 | 0.9998 | 0.9992 |
| 33 | 0.5496 | 0.3681 | 1.0000 | 1.0000 | 0.6358 | 0.4873 | 1.0000 | 0.9998 |
| 34 | 0.5384 | 0.3417 | 1.0000 | 1.0000 | 0.6297 | 0.4673 | 1.0000 | 0.9999 |
| 35 | 0.5272 | 0.3153 | 1.0000 | 1.0000 | 0.6236 | 0.4474 | 1.0000 | 1.0000 |
| 36 | 0.5160 | 0.2889 | 1.0000 | 1.0000 | 0.6174 | 0.4275 | 1.0000 | 1.0000 |
| 37 | 0.5048 | 0.2625 | 1.0000 | 1.0000 | 0.6113 | 0.4076 | 1.0000 | 1.0000 |
| 38 | 0.4936 | 0.2362 | 1.0000 | 1.0000 | 0.6052 | 0.3877 | 1.0000 | 1.0000 |
| 39 | 0.4824 | 0.2098 | 1.0000 | 1.0000 | 0.5991 | 0.3678 | 1.0000 | 1.0000 |
| 40 | 0.4712 | 0.1834 | 1.0000 | 1.0000 | 0.5929 | 0.3479 | 1.0000 | 1.0000 |