机场出租车问题研究

参赛队编号： 007

赛题类型代码： C

机场出租车问题研究

摘要：

如今，飞机是大众出行的主要交通方式之一，而出租车也是一个热门行业，因此送客去机场会是很多出租车司机面临的工作线路，出租车司机将乘客送到机场后面临两个选择，通过建模为出租车司机解决相关问题。

针对问题一，从出租车司机估计排队等候载客的时间与预期时间的比较来确定司机的决策方案。在出租车司机面临两个选择时，从时间角度来考虑，如若出租车司机估计排队等候载客的时间大于其返回市区拉客的时间，排队候车在消耗成本，而返回市区节省时间，会在市区内拉客赚钱，则明显出租车司机选择返回市区合理；如若出租车司机估计排队等候载客的时间小于其返回市区拉客的时间，返回市区时间减去出租车司机等待载客的时间，在剩下的时间中选择候车的司机正在拉客赚钱，而选择返回的司机正在消耗成本，则明显出租车司机选择排队等候载客合理。

针对问题二，我们查阅资料，以新郑机场为例，在误差允许的范围内，确定新郑机场一天 24 小时内每小时平均到港航班数量及其所占日流量比例，并统计出租车每小时内的平均收益，将这些数据带入模型验证模型的可行性与合理性。取月份和时间点这两个变量，分别跟蓄车池可接受的最大出租车数量求皮尔森相关系数，分别得到相关系数为 0.0295 和 0.4072，即说明相比于月份而言，时间点对该模型更具依赖性。

针对问题三，机场“乘车区”现有两条并行车道，当并行车道均建立出租车站点时，离出口远的出租车必须等到前方出租车接到客人驶出后才能驶出，这样可能会造成时间资源的浪费以及交通障碍；而一条行车道旁设置站点，另一条行车道供交通使用，即出租车接到客人后便可驶出，不必考虑前方出租车是否接到乘客，这样很大程度上节省了时间，也尽可能避免了交通堵塞。因此在一侧建立出租车站点，通过变量之间的计算来确定可建立出租车站点的数量。

关键字： 皮尔森相关系数，时间点，月份，新郑机场数据

一 问题重述

如今，飞机是大众出行的主要交通方式之一，而出租车也是一个热门行业，因此送客去机场会是很多出租车司机面临的工作线路，而在将乘客送到机场后，出租车司机会面临两个选择：

(A) 前往到达区排队等待载客返回市区。国内多数机机场送客和接客通道是分开的，出租车必须到指定的“蓄车池”排队等候，依“先来后到”排队进场载客。

(B) 直接放空返回市区拉客。

请结合实际情况，建立数学模型研究下列问题：

(1) 影响出租车司机决策的因素诸多，分析研究其主要因素的影响机理，综合考虑机场乘客数量的变化规律以及出租车司机的收益状况，为出租车司机建立选择决策模型，为其提供相应的选择策略。

(2) 针对国内某一机场及其所在城市的全部出租车，收集其相关数据，为出租车司机在该机场提供选择方案，并对问题一中的模型进行合理性检验，确定出租车司机对影响决策因素的依赖性。

(3) 一般会出现出租车排队载客和乘客排队乘车的情况。某机场“乘车区”现有两条并行车道，请给管理部门提供一种设置“上车点”，合理安排出租车与乘客，在保证车辆和乘客安全的条件下，使得总的乘车效率最高的方案。

(4) 机场出租车载客的收益与载客所需的行驶里程有关，乘客的目的地有远近之分，出租车司机不可选择乘客，也不能拒载，但允许出租车多次往返载客。管理部门拟对某些短途载客再次返回的出租车给予一定的“优先权”，请给出一种合理可行的“优先”方案，使得这些出租车的收益尽量均衡。

二 问题分析

2.1 问题一的分析

分析问题一，需要找出影响出租车司机决策的因素，从出租车司机估计排队等候载客的时间与预期时间的比较来确定司机的决策方案。在出租车司机面临两个选择时，应该从时间角度来考虑，如若出租车司机估计排队等候载客的时间大于其返回市区拉客的时间，即相同的时间内，排队候车在消耗成本，而返回市区节省时间，会在市区内拉客赚钱，则明显出租车司机选择返回市区合理；如若出租车司机估计排队等候载客的时间小于其返回市区拉客的时间，返回市区时间减去出租车司机等待载客的时间，在剩下的

时间中选择候车的司机正在拉客赚钱，而选择返回的司机正在消耗成本，则明显出租车司机选择排队等候载客合理。

2.2 问题二的分析

分析问题二，需要搜集数据来验证问题一所建立的模型，搜集的数据是问题一建立的模型的变量。我们查阅资料，以新郑机场为例，在误差允许的范围内，确定新郑机场一天 24 小时内每小时平均到港航班数量及其所占日流量比例，并统计出租车每小时内的平均收益，进行相关计算得到问题一所建立的模型的相关变量对应的数据，将这些数据带入模型验证模型的可行性与合理性。取月份和时间点这两个变量，求皮尔森相关系数[1]，根据最终结果来判断问题一中的模型对相关数据的依赖性。

2.3 问题三的分析

分析问题三，机场“乘车区”现有两条并行车道，当并行车道均建立出租车站点时，离出口远的出租车必须等到前方出租车接到客人驶出后，它才能驶出，这样可能会造成时间资源的浪费以及交通障碍；而在一条行车道旁设置站点，另一条行车道供交通使用，即出租车接到客人后便可驶出，不必考虑前方出租车是否接到乘客，这样很大程度上节省了时间，也尽可能避免了交通堵塞。因此在一侧建立出租车站点，通过变量之间的计算来确定可建立出租车站点的数量。

三 模型假设

1.假设乘客在上出租车后直接被送至目的地，不因某原因返程或半路下车。

2.假设天气，路况等因素均良好，对出租车司机载客无不良影响。

3.假设每辆出租车性能良好，差异较小可忽略不计。

4.假设返回市区之后两种方式的出租车获得的收益相同，且与时间相关。

5.假设出租车拉客与空车时行驶速度相同。

6.忽略人与人间的差异，假设他们一切都一样。

四 符号说明

符号	符号说明	符号	符号说明
W_1	出租车选择去蓄车池的	W_2	出租车选择返回市区的

	利润		利润
μ	选用出租车乘客的比例	λ	乘坐同辆出租车乘客的比例
n_{ij}	第 j 月第 i 时出租车的乘客人数	m	蓄车池现有出租车数量
T_1	出租车在蓄车池的等待时间	T_2	出租车空载返回市区的时间
T_0	再次进入蓄车池的出租车需要等待时间	W_3	出租车返回市区的收益
$\Omega^{'}$	拉客后出租车平均每分钟的收益	Ω_i	出租车拉乘客第 i 时内每分钟的收益
t	在将第一批乘客送往目的地之后，剩余时间	$\mu\lambda n_{ij}$	每个小时需要的出租车数量

五 模型的建立与求解

5.1 问题一的模型建立与求解

5.1.1 出租车司机在将乘客送到机场后，会面临两个选择：

（A）前往到达区排队等待载客返回市区。

（B）直接放空返回市区拉客。

记出租车选择去蓄车池的利润 W_1，出租车选择返回市区的利润 W_2，出租车在蓄车池的等待时间为 T_1 分，空载返回市区的所需时间为 T_2 分。

5.1.2 记选用出租车的乘客所占比例为 μ，同时坐一辆出租车的乘客占据比例 λ，例如如果两个人同时乘坐一辆出租车，那对他们各自而言 $\lambda = 0.5$.

5.1.3 记第 j 月第 i 时出租车的乘客人数为 n_{ij}（i =0, 1, ⋯, 23, j =1, 2, ⋯, 12, 表示第 j 月第 i 时内抵达该空港的乘客数量）。设每个小时需要出租车数量是 $\mu\lambda n_{ij}$，蓄车池内现在有 m 辆出租车，记再次进入蓄车池的出租车需要等待时间为 T_0 分. 返回市区的所需时

间为T_2分，忽略一些影响时间的相关因素，假设出租车返回市区所需时间是一个定值，则这段时间内的收益也是一个定值。设返回市区的收益为W_3，则

$$W_3 = T_2 \times \Omega^{'}$$

其中$\Omega^{'}$是一个定值，表示的是拉客后出租车平均每分钟的收益。

5.1.4 假设返回市区之后两种方式的出租车获得的收益相同，且与时间相关。设出租车拉乘客第i时内每分钟的收益为Ω_i（i=0, 1, …23, 表示该小时内出租车司机的平均收益)，在返回市区后，两种方式出租车在将第一批乘客送往目的地之后，剩余的时间为t.

5.1.5 比较数值，选择方案

目标函数：大致比较T_1和T_0数值大小。

如果T_1较大，表示现在的蓄车池需等待的时间小于预期时间，选择方案一，出租车进入蓄车池；如果T_0较大，表示蓄车池需等待的时间大于预期时间，选择方案二，出租车返回市区。

$$s.t.\begin{cases} W_3 = T_2 \times \Omega^{'} \\ T_1 = t \\ W_3 = t \times \Omega_i \\ T_0 = \dfrac{60m}{\mu\lambda n_{ij}} \end{cases}$$

5.2 问题二的模型建立与求解

5.2.1 查阅资料找到新郑机场附近出租车每小时的收益以及机场每小时平均到港航班数量及其所占日流量比例[2]。

新郑机场附近出租车每小时的收益图：

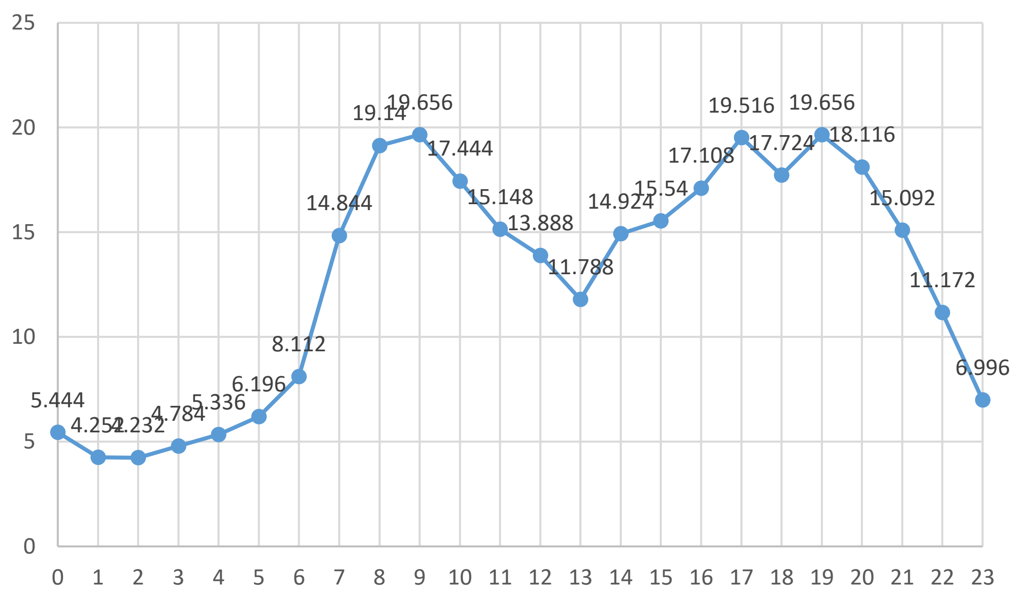

图 1 出租车每小时收益

新郑机场每小时到港航班情况：

表 1 新郑机场每小时航班情况表

时间	到港航班数量	所占日流量比例
0	278	3.34%
1	138	1.66%
2	32	0.39%
3	22	0.26%
4	10	0.12%
5	10	0.12%
6	7	0.08%
7	5	0.06%
8	478	5.75%
9	626	7.53%
10	416	5.01%
11	498	5.99%
12	524	6.30%
13	580	6.98%
14	588	7.07%

15	465	5.59%
16	479	5.76%
17	402	4.84%
18	500	6.02%
19	532	6.40%
20	552	6.64%
21	531	6.39%
22	322	3.87%
23	316	3.80%

以折线图的形式可以清楚地表示一天内新郑机场航班的变化情况：

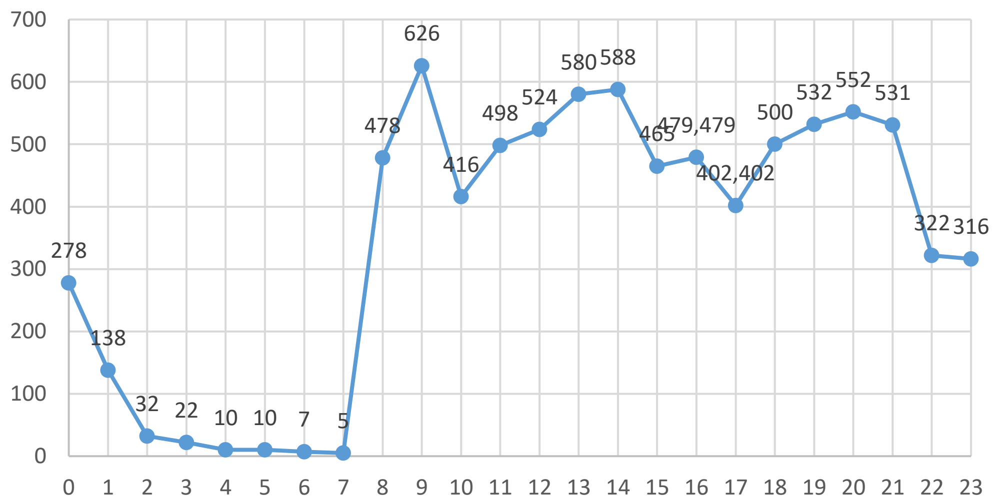

图 2 新郑机场每小时到港航班数量

5.2.2 在程序中需要新郑机场的月客运量的相关数据，但只查到了新郑机场的年客运量以及 2019 年我国民航月客运量的相关数据[2]，在误差允许的范围内，利用一定的比例系数，借助 2019 年我国民航月客运量的相关数据得到新郑机场的月客运量的相关数据。记比例系数为 k

$$k=\frac{\text{新郑机场年客运量}}{\text{2019年我国民航年客运量}}$$

2019 年我国民航月客运量的相关数据：

表 2 2019 年我国民航月客运量情况表

	国内民航客运量/万人
一月	4733.4
二月	4785.4
三月	4745.4
四月	4703.8
五月	4847.8
六月	4733.3
七月	5271.4
八月	5422.7
九月	4869.4
十月	5093.8
十一月	4715.4
十二月	4643.6
总计	58565.4

而经计算得

$$k = \frac{2912.93}{58565.4}$$

可得新郑机场月客运量的相关数据：

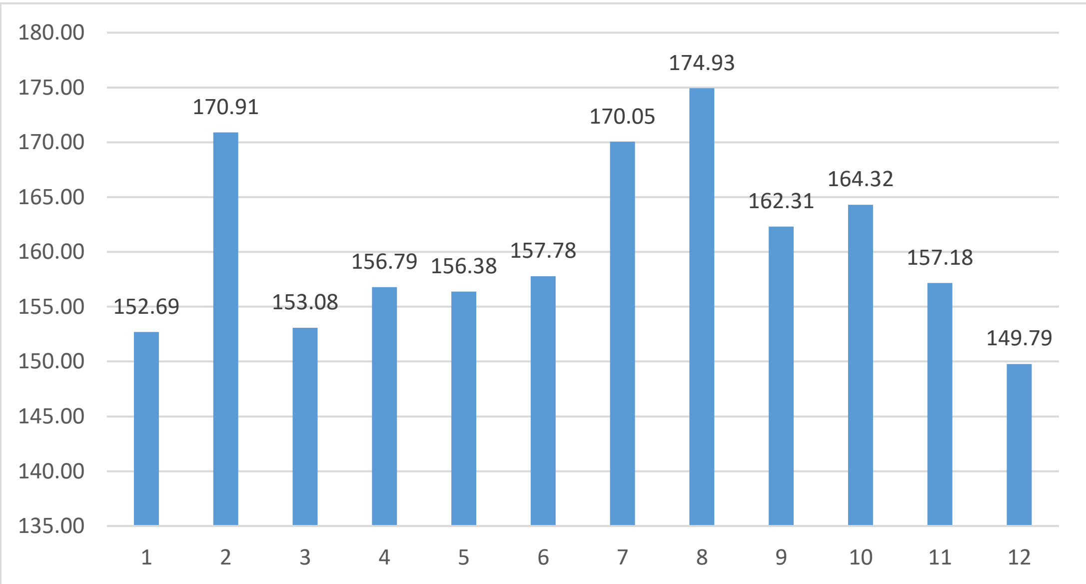

图 3 新郑机场月客运量的相关数据

5.2.3 设新郑机场每小时到达的人数是 x_i，此时的月份为 y_j，$n_{ij} = x_i \times y_j$。取随机抽

取某天到新郑机场获取数据，

$$
\begin{aligned}
&n_{ij}=x_i\times y_j\\
&\mu=0.15\\
&\lambda=\frac{1}{1.5}=0.667\\
&T_2=45\min\\
&\Omega^{'}=1.417
\end{aligned}
$$

对于出租车司机来说，最为直观的数据是当前蓄车池内出租车的数量，当时的时间以及月份。以这些数据为基础，可先求得在第i时、第j月内，可以接受的蓄车池最大出租车数量m_0，且$T_0=T_1$可求解：

$$m_0=\frac{45\mu\lambda\Omega^{'}}{60\Omega^{'}}$$

再将m_0与m输入的m进行比较，可得到该m值能否被出租车司机所接受，以此来判断选用方案一或方案二。最后计算选用该方案所得收入：若选用方案一，则收入为$T_2\times\Omega^{'}$，由于此时出租车已经承载了乘客，且返回市内使用高速公路，因此假设该值为一定值，$\Omega^{'}$的值由出租车司机每小时的最高收入决定；若选用方案二，则收入为$t\times\Omega_i$，其中Ω_i的具体数值在表格：出租车每小时收益中。整体程序于附录中的E_1.m.

在此研究的基础上，只考虑以上几个相关因素，并假设了当出租车返回市区后，即认为乘客已送至目的地。根据上述数据可以得出问题一中所建立的模型是合理的。

5.2.4 验证该模型的合理性

通过 *matlab* 运行程序，其参数设置为：60,7,10。表示当前蓄车池有 60 辆车，此时为 10 月某一天的 7 点。求解得到的m_0为：22，这表示司机能接受的出租车数量是 22 辆，但是现在有 60 辆，显然应选择方案二。并给出了选择方案二的收入，约为 181.63。再次运行，并将参数设置为 30,6,5。求解得到的m_0为 49，显然司机能接受的出租车数量大于现在蓄车池内出租车的数量，选择方案一，且收入约为 63.77。

5.2.5 取月份和时间点这两个变量，求皮尔森相关系数，根据最终结果来判断问题一中的模型对相关数据的依赖性。求皮尔森相关系数[1]的方法如下所示：

两个变量之间的皮尔逊相关系数定义为两个变量之间的协方差和标准差的商：

$$\rho_{X,Y}=\frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}=\frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y},$$

上式定义了总体相关系数，常用希腊小写字母ρ作为代表符号。估计样本的协方差和标准差，可得到皮尔逊相关系数，常用英文字母r代表：

$$r=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}.$$

r亦可由(X_i,Y_i)样本点的标准分数均值估计，得到与上式等价的表达式：

$$r=\frac{1}{n-1}\sum_{i=1}^{n}(\frac{X_i-\overline{X}}{\sigma_X})(\frac{Y_i-\overline{Y}}{\sigma_Y}).$$

其中$\frac{X_i-\overline{X}}{\sigma_X}$、$\overline{X}$及$\sigma_X$分别是对$X_i$样本的标准分数、样本平均值和样本标准差。

由于有两个自变量，但相关系数只能是一个自变量与因变量的结果，因此在运行时输入两个参数，表示当月份确定时，讨论时间与m_0的相关系数，反之当时间确定时，讨论月份与m_0的相关系数。经验证，当取$i=18$，$j=10$时运行程序 E_2.m，得到i与m_0的相关系数为 0.4072，j与m_0的相关系数为 0.0295，其散点图如下所示：

好是需要一个递增或递减的趋势，因此出现较大误差。

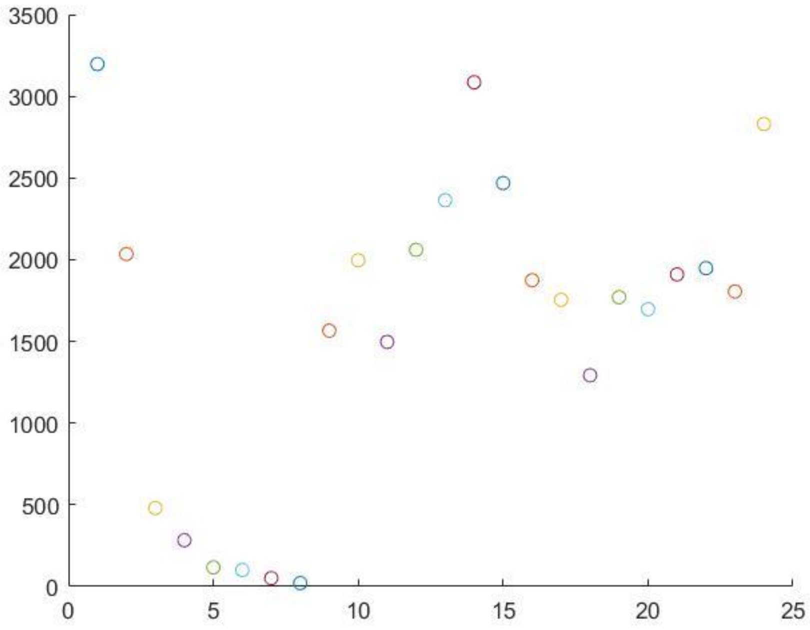

图 4 i 与 m_0 的关系图

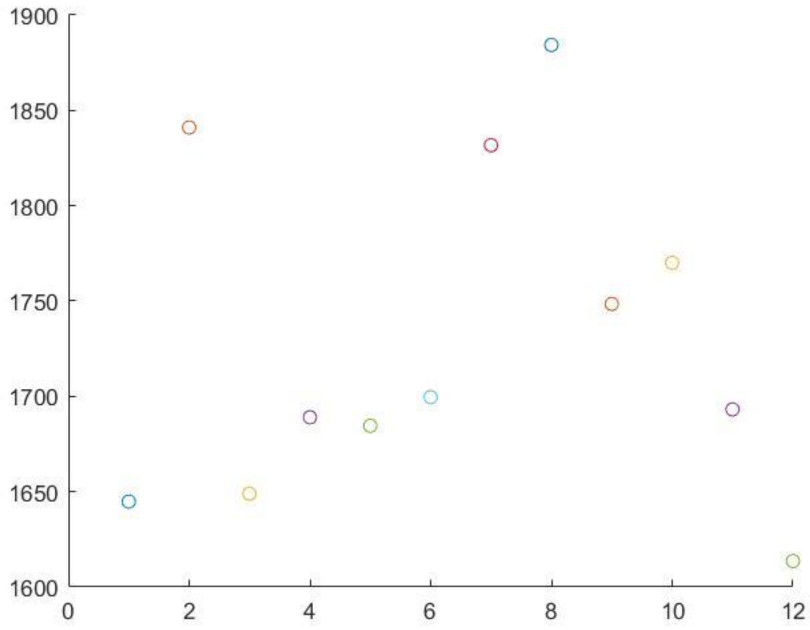

图 5 j 与 m_0 的关系图

经分析，相关系数较低的主要原因是：时间和月份是连续分布的，而相关系数最好是需要一个递增或递减的趋势，因此出现较大误差。

5.3 问题三的模型建立与求解

5.3.1 机场“乘车区”现有两条并行车道，当并行车道均建立出租车站点时，离出

口远的出租车必须等到前方出租车接到客人驶出后，它才能驶出，这样可能会造成时间资源的浪费以及交通障碍；而在一条行车道旁设置站点，另一条行车道供交通使用，即出租车接到客人后便可驶出，不必考虑前方出租车是否接到乘客，这样很大程度上节省了时间，也尽可能避免了交通堵塞。

需乘坐出租车的客人在候车区排队，由于成员人数不同，记有a的乘客需要单人打车，b的乘客两人一起打车，c的乘客三人或四人一起打车，其中：

$$a+b+c=1.$$

而一个人上车（包括装行李，等人等行为）需要时长d，新的出租车到达空出的站点需要时长为e，每个站点之间的距离为均值记为l，乘客的步行速度记为s，则每个站点接到乘客需要等待时间为I：

$$I=(a+2b+4c)\times d+e$$

站点围绕中心两侧分布，则一侧可建立n个站点（n必为整数），即等到中心位置的车站可以再次等待乘客时，新的乘客可以走到离中心第n个站点搭乘出租车。

按照这样推算，整体就需要建立N个出租车站点，

$$N=2n+1$$

这样就可以使得每组乘客都有空站点选择乘坐出租车，满足题目要求。

5.3.2 代入收集的数据检验

（1）假定乘客排队而来，其中有 50%的乘客需要单人打车，20%的乘客两人一起打车，30%的乘客三人或四人一起打车。

（2）1 个人上车平均需要 20 秒，新的出租车到达站点需要 10 秒，每个站点之间的距离均为 6 米，乘客步行的速度均为每秒 1 米

（3）代入上式中可得到

$$I=1*20*0.5+2*20*0.2+4*20*0.3+10=52\text{秒}$$

即平均每个站点接到人需要等 52 秒。

$$n=\frac{52\times1}{6}=9$$

站点围绕中心两侧分布，则一侧可建立 9 个站点即等到中心位置的车站可以再次等待乘客时，新的乘客可以走到离中心第 9 个站点搭乘出租车。

于是最多只需要建设$1+2\times9=19$个站点就可以满足每组乘客都有空站点可以达成出租车。

六 模型的评价

6.1 模型的优点

（1）本文在正确、清楚地分析了题意的基础上，建立了合理、科学的数学模型，该模型能够将一系列大数据进行处理，从而得到可靠结果。

（2）模型较为符合现实，便于检验，易于应用。

（3）模型的计算采用了专业的数学软件，可信度较高。

（4）对模型中相关因素进行了量化分析，查阅了大量的实际且具体的数据，使论文具有说服力和真实性。

6.2 模型的缺点

（1）模型中的一些变量对模型的依赖性比较强，而部分数据查询不够全面精准，使其存在一定的误差。

（2）考虑的影响因素较少，在处理问题时可能存在一些误差。

七 模型的推广与改进

7.1 模型的推广

本文的模型除了适用于机场的出租车问题，还适用于客运站拉客等一系列接客问题。

7.2 模型的改进

7.2.1 本文主要考虑出租车司机在两种方案内各自所需要的时间，但影响出租车司机决策的因素还有很多，比如天气、意愿等，还可以将更多因素考虑进去对模型进行改进。

7.2.2 由于模型中的一些变量对模型的依赖性比较强，对于数据的掌握必须精准且全面，这样会减少误差。

八 参考文献

[1]陈功平、王红，《改进 Pearson 相关系数的个性化推荐算法》，山东农业大学学报（自然大学学报），2016

[2]国家统计局的数据统计，百度百科

附录：

程序 E_1.m：

```
function chooseplan(m,i,j)
%CHOOSEPLAN 该函数用于计算在 j 月第 i 时，当出租车到达航站楼，蓄车池内有 m 辆车时，对两种方案进行决策

% 读取第 i 时新郑机场每小时航班到达比，获取 xi
flight=xlsread('新郑机场每小时航班到达比.xlsx');
x=flight(i+1,3);



% 读取 j 月份平均每天全国客运量
passenger=xlsread('2019 年我国民航月客运量.xlsx');
y=passenger(j,2)*(2912.93/58565.4)*10000;



% 读取第 i 时司机平均收入
income=xlsread('出租车每小时收益.xlsx');
omiga=income(i+1,2)/60;
Omiga=1.417;





n=x*y;
miu=0.15;
lamuda=1/1.5;
```

```
% 获取可以接受的最大m0
m0=45*Omiga*miu*lamuda*n/(60*omiga);
disp('方案一可以接受的蓄车池出租车数量');
disp(m0);

if m<m0
    disp('应选择方案一');
    W=45*Omiga;
    disp(W);
else
    disp('应选择方案二');
    W=60*m/(miu*lamuda*n)*omiga;
    disp(W);
end

end
```

程序E_2.m：

```
function [coeff1,coeff2]=pearson(i,j)
%PEARSON 求皮尔森相关系数



% 读取第i时新郑机场每小时航班到达比，获取xi
flight=xlsread('新郑机场每小时航班到达比.xlsx');
x=flight(i+1,3);
parameter1=flight(:,3);
parameter1=reshape(parameter1,1,24);
```

```
% 读取 j 月份平均每天全国客运量
passenger=xlsread('2019 年我国民航月客运量.xlsx');
y=passenger(j,2)*(2912.93/58565.4)*10000;
parameter2=passenger(:,2)*(2912.93/58565.4)*10000;
parameter2=reshape(parameter2,1,12);


% 读取第 i 时司机平均收入
income=xlsread('出租车每小时收益.xlsx');
omiga=income(i+1,2)/60;
parameter3=income(:,2)/60;
parameter3=reshape(parameter3,1,24);

Omiga=1.417;


res=(1:24);
time=(1:24);
for t=1:24
    n=parameter1(1,t)*y;
    miu=0.15;
    lamuda=1/1.5;

    res(1,t)=45*Omiga*miu*lamuda*n/(60*parameter3(1,t));
    figure(1);
    scatter(time(1,t),res(1,t));
    hold on;
end
```

```
fenzi=sum(time.*res)-(sum(time)*sum(res))/length(time);
fenmu=sqrt((sum(time.^2)-sum(time)^2/length(time))*(sum(res.^2)-
sum(res)^2/length(time)));
coeff1=fenzi/fenmu;




res2=(1:12);
month=(1:12);
for t=1:12
    n=x*parameter2(1,t);
    miu=0.15;
    lamuda=1/1.5;

    res2(1,t)=45*Omiga*miu*lamuda*n/(60*omiga);
    figure(2);
    scatter(month(1,t),res2(1,t));
    hold on;
end
fenzi=sum(month.*res2)-(sum(month)*sum(res2))/length(month);
fenmu=sqrt((sum(month.^2)-sum(month)^2/length(month))*(sum(res2.^2)-
sum(res2)^2/length(month)));
coeff2=fenzi/fenmu;



end
```