*Hollow Man*

# 机场出租车问题的最优策略


**摘　要：**

现代社会中，因为飞机这种交通工具的便捷性，越来越多的人选择坐飞机外出旅游。同时，又因为机场的飞机起降噪声大，同时机场所需飞机起降的空间较大，所以机场一般远离市区，因而许多乘客选择乘坐出租车到达市内目的地。在机场接客的出租车司机，为了使得收益最大化，需要权衡等待载客返回市区和直接返回市区拉客的收益。同时，机场管理者需要合理安排上车点，使得乘车效率最高。为了体现人文关怀，使得拉短途和长途的出租车司机的收益尽量均衡，机场管理者还需要一个短途载客再次返回的出租车"优先" 安排方案。

本文通过M/M/1 排队模型和 ARIMA 历史数据分析模型解决了出租车司机的选择问题，给出租车司机提供了一定的参考价值。同时运用 Monte Carlo 仿真和模拟遍历的方法，解决了乘车效率问题和均衡出租车司机收益问题，为机场管理者提供了一些合理的建议。



**关键词：　排队论，出租车，仿真模拟**

# 一、 问题重述

## 1.1 情况说明

大多数乘客在下飞机后需要离开机场，乘坐交通工具到达目的地。出租车是主要的交通工具之一。因为机场大多将出发与到达通道分开，所以送客到机场的出租车司机需要根据可观测到的确定信息即在某时间段抵达的航班数量和“蓄车池”里已有的车辆数两种可能性中做出选择：

(A) 前往到达区排队等待载客返回市区。出租车须要到达指定的“蓄车池”排队等候，依“先来后到”的原则排队进场载客。很明显，等待时间长短取决于排队出租车和乘客的数量多少。此方案需要付出一定的时间成本。

(B) 直接空载返回市区拉客。 选择此方案，出租车司机会付出空载费用和可能损失潜在的载客收益。

## 1.2 相关信息

通常司机的决策与其个人的经验判断有关，比如在某个季节与某时间段抵达航班的多少和可能乘客数量的多寡等。在实际中，还有很多影响出租车司机决策的确定和不确定因素，其关联关系各异，影响效果也不尽相同。

同时，如果乘客在下飞机后想“打车”，就要到指定的“乘车区”排队，按先后顺序乘车。机场出租车管理人员负责“分批定量”放行出租车进入“乘车区”，同时安排一定数量的乘客上车。

## 1.3 需要解决的问题

问题 1 要求分析研究与出租车司机决策相关因素的影响机理，同时综合考虑机场乘客数量的变化规律和出租车司机的收益， 建立出租车司机选择决策模型，并给出司机的选择策略。

问题 2 要求收集国内某一机场及其所在城市出租车的相关数据，根据问题 1 的出租车司机选择决策模型给出该机场出租车司机的选择方案，并分析模型的合理性和对相关因素的依赖性。

问题 3 需要规划出“上车点”，并合理安排出租车和乘客，在保证车辆和乘客安全的条件下，使得两条并行车道的总的乘车效率最高。

问题 4 要求在满足出租车司机不能选择乘客和拒载，但允许出租车多次往返载客的条件下，因为乘客的目的地有远有近，机场的出租车载客收益与载客的行驶里程有关，所以需要规划出一个可行的短途载客再次返回的出租车“优先” 安排方案，使得这些出租车的收益尽量均衡。

## 二、 问题分析

### 2.1 问题 1，2 的分析

如果选择方案 A，那么司机需要付出空载时油费的代价，但是节约了时间成本。如果选择方案 B，那么司机需要付出时间成本，但是载客返回市区会给予司机一定的利润。方案 A 的一次接客循环是空载返回市区，市区拉客到机场。方案 B 的一次接客循环是等待乘客上车，拉客返回市区，市区拉客到机场。因为这两种方案所消耗的时间不同，所以衡量这两种接客循环方案利润的指标是每分钟赚钱的数目。

### 2.2 问题 3 的分析

因为我们需要规划“上车点”泊位数量，所以首先我们将基于一种上车点泊位设计模型进行设计。又因为出租车上客区运行具有供给限制性、空间局限性、系统随机性[2]，很难用数学解析的方法建立模型求解出合理的出租车泊位数量，因此可以采用Monte Carlo 方法进行仿真分析，这样不仅可以获得系统的整体特征指标，同时还能对其中间的过程特性进行分析，有利于更好地理解系统的运行过程。

### 2.3 问题 4 的分析

此问题可以用程序模拟实现多辆出租车载客收益。并由程序遍历所有可能的划分，每种划分均需求出在此情况下多名司机一个月的收益，并对多名司机的受益求取方差，比较所有可能的方差，并寻找其中方差最小的划分情况，重复多次，求取多次的最优划分的数据，并求对应的期望作为理想的划分情况。

## 三、 问题假设

为了简化问题，我们首先做出几条总体的基本假设：

（1）在机场上车点，所有乘客在乘车点排队上车，且服务情况为先到先得。

（2） 1 辆出租车仅搭载 1 位乘客。

（3）乘客候车区容量无限，且出租车是固定需求的，即乘客不会因为排队等待时间过长而离开，队列不会发生中途流失。

（4）在不重叠时间区间内乘客到达数是相互独立的。

（5）不考虑出租车和乘客的移动时间。

（6）每班航班到达的人数相同，出租车和乘客到达上客区的时间服从负指数分布。

为了简化模型，关注问题本身，所以我们又分别对各题做出以下假设：

### 3.1 问题 1,2 的假设

（1）忽视出租车到市区等待接客的这一时间，即假设出租车到市区立马能接到客。

（2）所有乘客到市区下车的地方与机场为一固定距离。

### 3.2 问题 3 的假设

假设蓄车池的容量无限大，可以让所有想进来的出租车都可进入。

### 3.3 问题 4 的假设

假设我们可以开辟多个蓄车池，设为蓄车池 1，2，3。并依照出租车上次载客路程，将出租车下一次的载客优先级分为 1，2，3，三级。

三级优先度可在 1，2，3 蓄车池内任意挑选一条车少的蓄车池，并进入其中等待乘客，二级优先度可在 1，2 蓄车池内任意挑选一条车少的蓄车池，并进入其中等待乘客，一级优先度只能选择 1 号蓄车池。

## 四、 模型准备

我们选取了互联网上航空公司官网所公开的一些数据，并且通过数据处理的形式，

转化成为软件可利用的形式，从而便于进行数据分析。

# 五、 模型选取、求解、分析

## 6.1 问题 1

首先来考虑乘客排队队列中的人数。

由题意，令 N(t)表示在(0，t]时间内到达的乘客数。 $P_n(t_1,t_2)$表示在时间段$[t_1,t_2]$内有 n 名乘客到达的概率。假设 $P_n(t_1,t_2)=P\{N(t_1)-N(t_2)=n\}$，那么在充分小的△t 时间，在时间区间[t, t+△t)内有一个乘客到达的概率与 t 无关，为了简化计算，我们认为其与△t 成正比，即：P1(t, t+△t)=λ△t +o(t)。λ表示单位时间内平均到达的乘客数。对于充分小的△t，在时间区间内有两个或两个以上的乘客到达的概率极小，可以忽略，即$\sum_{n=2}^{2} P_n(t, t+\triangle t) = o(\triangle t)$，整个乘客的进入过程服从泊松分布。那么在充分小的△t 时间，在时间区间[t, t+△t)内有一个乘客上车的概率与 t 无关，而大约和区间长△t 成正比，即 P（T≤△t）=μ△t +o(△t)，T 为服务时间，μ表示单位时间内平均服务的乘客数。由于假设只有一个上车点，所以多人同时完成服务的概率为零。则在 t+△t 时刻的排队人数由 t 时刻转化而来，共有三种情况：

1. t 时刻一共有 n 名乘客在排队，且在△t 时间内没有乘客到达且没有乘客服务完成，概率为 1-λ△t-μ△t +o(△t)。

2. t 时刻一共有 n+1 名乘客在排队，且在△t 时间内没有乘客到达且有一乘客服务完成，概率为μ△t +o(△t)。

3. t 时刻一共有 n-1 乘客在排队，且在△t 时间内有一乘客到达且没有乘客服务完成，概率为μ△t +o(△t)。

令 $P_n(0, t+\triangle t)$为 t+△t 时刻有 n 名乘客的概率。则：

$P_n(0, t+\triangle t)$

$= P_n(0, t)(1-\lambda\triangle t-\mu\triangle t)+ P_{n+1}(0, t)\mu\triangle t+ P_{n-1}(0, t)\lambda\triangle t+ o(\triangle t)$

$= \frac{P_n(0,t+\triangle t)-P_n(0,t)}{\triangle t}$

$= \lambda P_{n-1}(0, t)+\mu P_{n+1}(0, t)-(\lambda+\mu)P_n(0, t)+\frac{o(\triangle t)}{\triangle t}$

令△t → 0，则$\frac{dP_n(0,t)}{\triangle t}=\lambda P_{n-1}(0, t)+\mu P_{n+1}(0, t)-(\lambda+\mu)P_n(0, t)$

当 n=0 时，$P_n(0, t+\triangle t)=P_0(0, t)(1-x\triangle t)+P_1(0, t)(1-\lambda\triangle t)\mu\triangle t+ o(\triangle t)$

令△t → 0，则 $\frac{dP_0(0,t)}{dt}=-\lambda P_0(0, t)+\mu P_1(0, t)$

因为概率与 t 无关，所以导数为 0

$$\begin{cases} \lambda P_{n-1}(0,t)+\mu P_{n+1}(0,t)-(\lambda+\mu)P_n(0,t)=0 \\ -\lambda P_0(0,t)+\mu P_1(0,t)=0 \end{cases}$$

化简得：

$P_n(0, t)= (\frac{\lambda}{\mu})^n P_0(0, t)$

因为$\sum_{n=0}^{\infty} P_n(0,t)=1$（概率性质），$P_0(0, t)=1-\frac{\lambda}{\mu}$

所以 $l_s=\sum_{n=0}^{\infty} nP_n(0,t)=\sum_{n=0}^{\infty} n\left(1-\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n=\frac{\lambda}{\mu-\lambda}$

其中λ和μ的取定与到达的航班数（进入排队队列的人数）和前面等待的出租车人数（出排队队列的人数）有很大的关系，对于这两个重要参数，我们使用 ARIMA 模型利用所搜寻到的历史数据来对之后的情况进行预测，从而求出λ和μ并对出租车的决策进行决定。ARIMA 模型全名为差分自回归移动平均模型[1]，它是由自回归模型（AR）与移动平均模型（MA）结合所形成的 ARMA 模型的延伸，公式如下：

$$y_t=\mu+\sum_{i=1}^{p}\gamma_i y_{t-i}+\varepsilon_t+\sum_{i=1}^{q}\theta_i\varepsilon_{t-i}$$

使用 AR 模型，MA 模型和 ARMA 模型要求数据是平稳的，即要求经由样本时间序列所得到的拟合曲线，在未来一段时间内仍能顺着现有状态延续下去，但是这与事实并不相符。我们可以使用差分法对波动很大的数据进行平稳，即时间序列在 t 时刻与 t-1 时刻的差值。一次即为一阶差分，对于一阶差分后再做差分就能到得到二阶差分。如果在对于数据进行 d 阶差分以后能够转换成一个平稳的时间序列，那么就可以对差分后的数据建立 ARMA 模型。ARIMA (p, d, q)模型中 AR 是自回归，p 为自回归项；MA 为移动平均，q 为移动平均项数，d 为时间序列成为平稳时所做的差分次数。

具体步骤如下:

1. 将搜集到的数据导入到程序之中。

2. 对于输入进的数据进行 d 阶差分计算，化为平均时间序列。

3. 对于平均时间序列分别求得其自相关系数 ACF 和偏自相关系数 PACF，通过对自相关图和偏自相关图的分析，得到最佳的阶层 p 和阶数 q。

4. 由以上得到 d, q, p 得到 ARIMA 模型。

相关性是有序的随机变量序列与其自身相比较，自相关函数反映了自身数据在同一序列在不同时间之间的相关性，公式为:

$$ACF(k) = \rho_k = \frac{Cov(y_t, y_{t-k})}{Var(y_t)}$$

自相关函数 ACF 的取值范围为[-1,1]，-1 为负相关，+1 为正相关，0 为无相关性，一般会取 95%的置信区间。

ACF 还包含了其他变量的影响，而偏自相关系数 PACF 是严格这两个变量之间的相关性。

假设每架飞机有 k 人，可计算出在某一时间段单位时间内所到达的人数，此时的单位时间内所服务的人数基本只与乘客数量的多少和行李多少有关，平均服务时间为 m 秒，那么司机所等待的时间为平均对长除以评价服务时间，假设司机去市区的需要 t 分钟，路长 l，司机载客开车每公里能赚到 a 元，空载每公里损失 b 元。计算两种情况下司机每分钟所能赚到的钱并进行比较，可以确定司机应该选取哪种方案。

方案 A 每分钟可以获利：$\frac{(a-b)l}{2t_0}$

则方案 B 每分钟可以获利：$\frac{2al}{t+2t_0}$

对于出租车司机来说，假如每分钟赚钱一样，那么司机会选择时间少的那个方案。

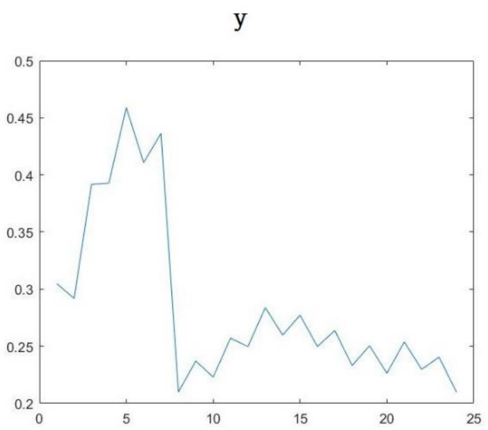
则当 $t_1 > \frac{4x_2t_0}{x_2-x_1} - 2t_0$时，选择 A 方案划算，否则选择 B 方案。

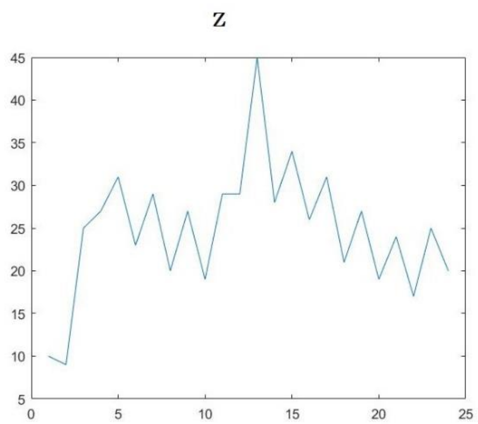
其中 $t_1$ 与排队队列中的人数有关。

### 6.2 问题 2

由网络了解到，按照天津的出租车价位，3km 以下，起步价 8 元，超过 3km 但不超过 9km 的部分，每千米收费 1.7 元，超过 9km 则每千米加收 2.25 元。目前 95 号汽油价格为每公升 7 元，正常出租车 100 公里需要消耗掉 8L 汽油，一共就是 56 元油费。平均到每 1 公里就是 0.56 元人民币。常见的空客 A320 飞机共有头等/公务舱 12 个座位，经济舱 173 个座位，总计 185 个座位。机场到市区常见距离为 15km, 出租车花费平均时间为 20min。乘客的平均上车时间为 10s。

通过搜集而来的数据，可知 l=15km，a=2.11 元，b=0.56 元，t=20min，k=185 人，m=10s，y 为该日 24 小时内每隔一小时时间段内预测每分钟航班数，z 为特定日期 24 小时内每隔一小时时间段内的预测出租车数量，y 和 z 都是由 Matlab 运行 arima.m 预测得到的，使用的数据已经包含在源代码中，预测如下图所示。




通过问题 1 中的模型，使用 Matlab 运行 Q2.m，得到以下结论：

| 时 间 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 选 择 方 案 | A | A | A | A | B | A | B | A | A | A | B | B | B | B | B | A | B | A | A | A | A | A | A | A |

对照上图和上表，当出租车等待数较多时，选择 B 方案的可能性较高，此符合常理，并且推理逻辑严密，因而此模型较为合理。对于收集而来的数据进行分析可知，天气，季节，不同时间段下飞机的人乘坐出租车的人数的比例都会影响到实际司机的选择。

## 6.3 问题 3

查阅资料，常见的出租车上客区设置模式可分为 3 种[3]：依次发车模式、单独发车模式和并列多车道发车模式。我们为了研究方便，仅考虑单车道设置上车点的情况，这种情况又分为依次发车模式和单独发车模式两种具体的情况。同时，又考虑到要充分利用两条并行车道的路况，并且为了安全，我们决定采用单独发车模式来进行上车点的构建。

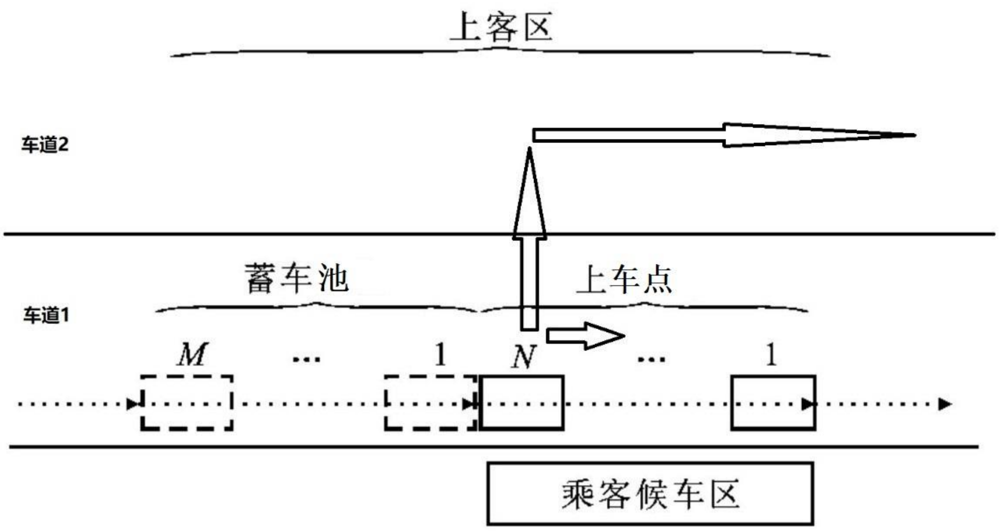



图 1  上客区乘客和出租车的设置

如图，在充分利用两条车道的情况下，我们采用以上设计模式，车道 1 乘客候车区左侧部分为蓄车池，供出租车车辆排队等候，车道 1 乘客候车区对应部分为上车点，乘客将有序乘上出租车。上客区包括M个蓄车泊位、N个上车点，上车点长度取值应该适应车身的大小和安全驶出车道距离。出租车载上了客人，司机可以选择变道到车道 2 或者沿着车道 1 继续行驶，来避免排队发车离去，最大化效率，此时出租车在上车点的停留时间仅受乘客上车时间的影响。

下面我们采用Monte Carlo 方法进行仿真分析，其具体步骤如下：

(1) 仿真初始化。给定出租车上车点的泊位规模 M，出租车、乘客到达时间间隔 tv，tp、乘客上车时间的随机分布参数和系统最大仿真时间，初始化仿真时间 T = 0 （单位：秒）。生成随机数，从而确定出租车和乘客到达的仿真时间。

(2) 以 1 s 为增长步阶更新仿真时间 T 。

(3) 出租车和乘客的到达。在该仿真时间时判断是否应有出租车或乘客到达。

有出租车或乘客到达时，进入蓄车池或乘客候车区排队，并生成下一出租车或乘客到达的时刻。

(4) 出租车驶离。完成载客过程的出租车后，驶离上车点。

(5) 待客出租车重排。上客区的空载出租车进行重排。空载出租车依次向前驶离蓄车池，占用所有腾空的上车点。

(6) 乘客上车过程。根据乘客上车规则，候车区排队乘客优先选取车队前端的待客车辆，并随机生成上车时间从而确定该车辆的驶离仿真时间。

(7) 判断仿真结束条件。当仿真计时器 T 未达到最大仿真时刻时，返回(2)；否则，仿真结束。

参考[6]中的数据，我们决定取用 7200s 的最大仿真时间，且根据实际情况，设定每 90s 到达一辆出租车和一名乘客，乘客的平均上车时间为 10s。

我们根据以上步骤构建了 Python 程序，对一般机场可行的上车点个数（1-30）分别进行仿真模拟，每种可能上车点数量的乘车效率结果取 5 次平均值，运行 Q3.py，结果如 Q3.xlsx 表和下图所示。

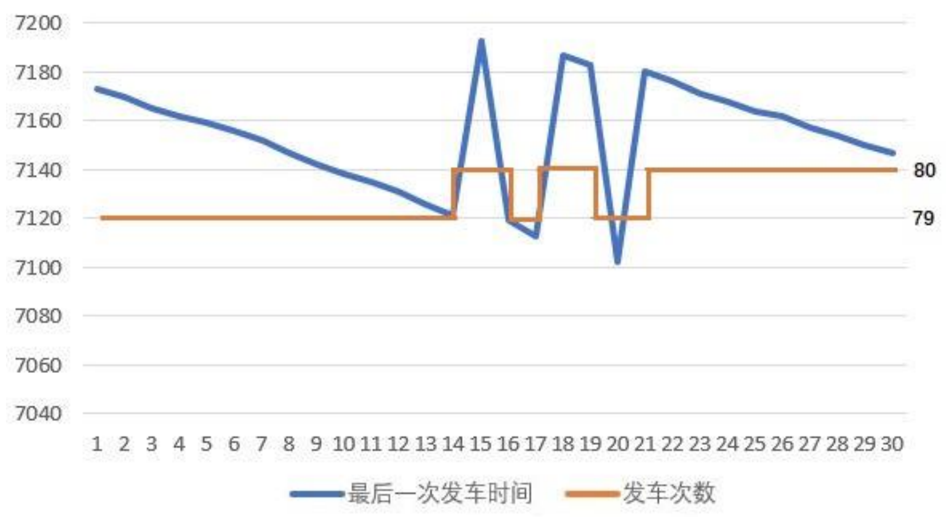



图 2 不同上车点乘车效率的比较

根据仿真结果，我们可以得出结论，以设定的上客区模型样式，上车点设置得越多，对于司机和乘客而言总的乘车效率也就越高，然而在实际应用中，由于场地的限制和为了确保乘客安全，上车点无法无限制的增加，我们应该选择合适的车间距来规划上车点，设定尽可能多的上车点，其效率最高。

## 6.4 问题 4

使用 Java 建立三个元素 taxi 类的队列作为三个蓄车池，建立 6 个不同的线程：aTaxi，bTaxi，cTaxi，dTaxi，eTaxi，fTaxi。假设机场到市区至少 15 公里，则每次再给定一个在 15 公里之上的路程范围，例如给定 30km，则实际路程范围为 15km~（15+30）km，每辆出租车都由随机数生成在此路程范围区间内的路程数，并计算收益，当出租车再次回到蓄车池时，根据自己的优先级和蓄车池内车辆的多少自动选择应进入的蓄车池，最后由 6 辆出租车在 30 天内的总收益计算方差。并自动显示出最优的划分方案（方差最小）。

使用 Java 运行程序 taxi.java 后，得出结果：在路程范围为 15 到 45 公里时最合理的优先级划分为：三级：0-22km，二级：22-33km，一级：33-45km。

如果是其它不同的城市可更改市区到机场的距离，以及对应的出租车起步价等信息，即可自动模拟不同城市的最佳优先级划分。

## 六、 模型的评价与推广

对于问题 1，2，使用M/M/1 排队模型和 ARIMA 历史数据分析模型对于出租车司机拉客选择的问题的建模较为精准，准确度高。对于问题 3，4，模拟仿真的实现较容易，并且易于获得较准确的解。

然而，在实现建模的过程中，因为时间较为紧张，且数据比较贫乏，我们没有考虑实际问题中的一些可能性来进行建模，而是将其理想化，这些遗憾包括：

（1）未考虑出租车可以搭载多位乘客的情况

（2） 1 辆出租车仅搭载 1 位乘客。

（3）乘客候车区容量并非是无限的，出租车并非为固定需求的，即乘客可能会因为排队等待时间过长而离开，而且许多乘客可能会选择机场大巴等便宜的交通工具返回市区。

（4）处于不同位置的出租车和乘客需要的移动时间不同，这些因素未考虑到问题 3 中。

（5）问题 1 中出租车到市区等待接客的这一时间未考虑在内，即出租车到市区

并非立马能接到客，而且所有乘客到市区下车的地方与机场是变化着的。

## 七、 参考文献

[1] 林思睿. 机场出租车运力需求预测技术研究[D].电子科技大学,2018.

[2] 黎冬平. 出租车上客区规模的 Monte Carlo 仿真计算方法[J]. 计算机工程与应用, 2011, 47(18): 22-25.

[3] 吴娇蓉, 李 铭, 梁丽娟. 综合客运枢纽出租车上客点管理模式和效率分析[J]. 交通信息与安全, 2012, 30(4): 18-23