问题二仿真中的假设：

假设位置因素、价格因素、距离因素对任务订单的完成影响是独立的。

## 5.2 问题二的模型建立与求解

问题二要求我们制定一个新的任务定价方案并与已经完成的旧的定价方案进行比较来说明当前定价方案的优点与缺点在何处。

方案的合理性主要从价格的分布以及任务的完成率两个方面来进行考量，任务的完成率我们采用蒙特卡洛模拟的方式对已完成订单进行模拟，经过100次模拟，订单的完成数目均值为521.9，完成率为62.5029% ，与给定完成订单的数据522及完成率62.5149%基本一致，故认为当前模拟方案可行。——-摘要

### 5.2.1 模型的准备

预定成功的概率

首先，会员订单的分配争抢过程是按照其信誉度来进行配额的，对次我们通过对会员预订任务开始的时间进行机理性分析可以得到所有会员的选单分31批次进行，每批次间隔3分钟，总共的会员数为1877人，除了最后一轮有1059人外，其他30轮的会员分布都比较分散，会员任务分配时是按照预计限额来配发，也就是说限额大的选中订单的有优先选择的概率，也就是限额大的预定订单成功概率大，我们可以用每个订单的限额与一定范围内且为当前轮次的所有订单限额的比值做为概率，其表达式大致如下



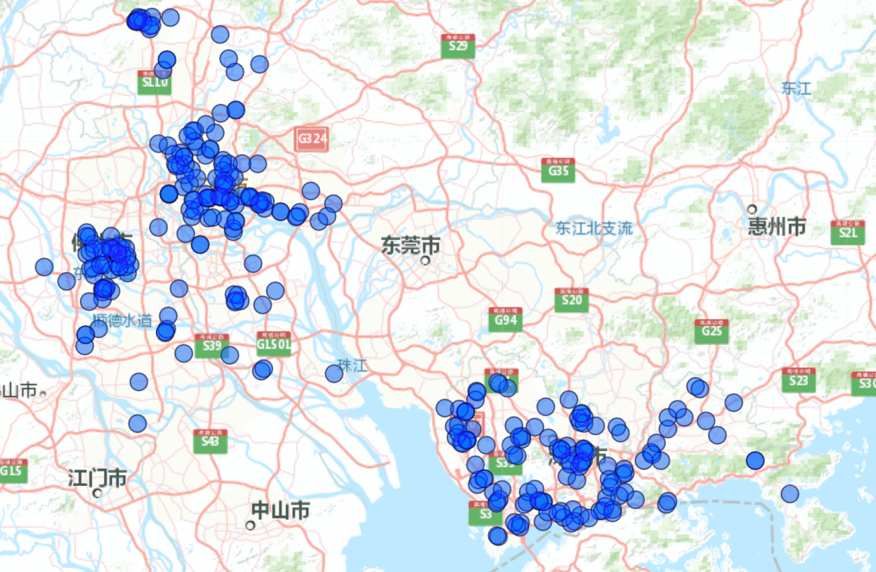
其中，表示当前订单被选择上的概率，

任务完成的概率加解释

其次对于订单的定价方案对于任务是否完成的影响，我们主要根据最后当前订单的完成率来衡量定价方案是否合理，我们首先考虑影响订单完成率的影响因素，我们大致确定了影响订单的主要因素主要有一下几个方面。

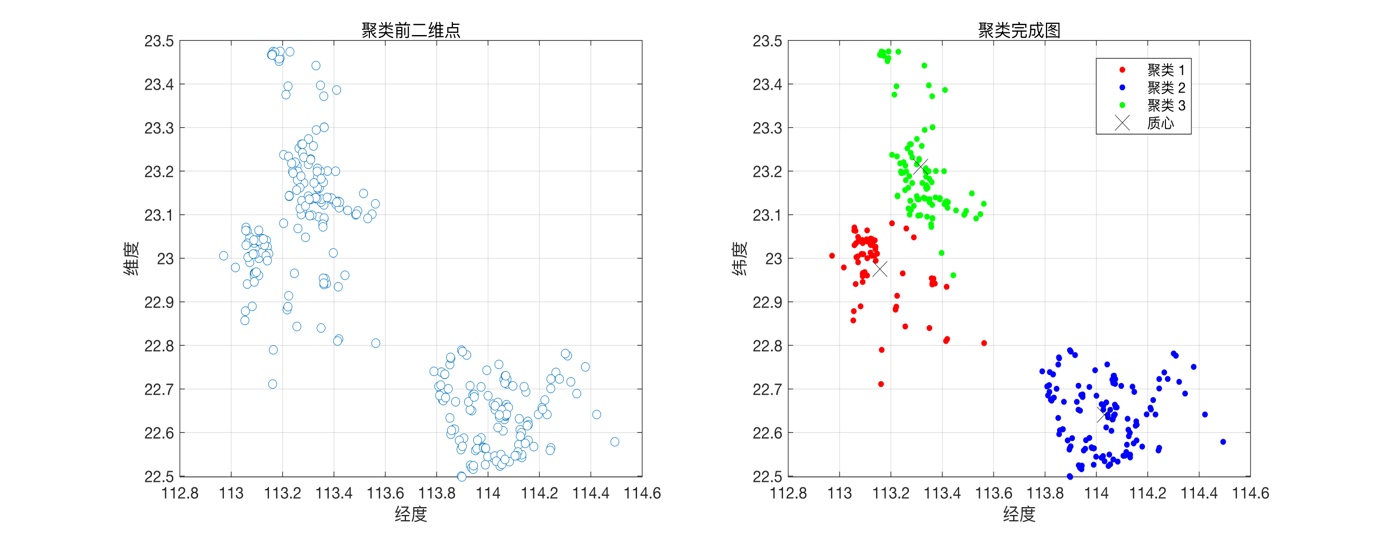
（一）位置因素

可视化未完成订单的位置分布，将所有未完成的订单映射到真实地图上，如下图



未完成订单的真实位置

通过直观观察，可以看出未完成的订单中的大部分集中在三个块中，为了更好的得出这种聚集因素，结合已有的完成订单数据通过k-means聚类将未完成的订单根据地理位置分作3类，聚类结果如下：



未完成订单聚类图

得到聚类后三个质心的经纬度分别为{113.1571 , 22.9753} , {114.0312 , 22.6409} , {113.3156 , 23.2101} 对应到真实的地图上分布，可以得到这三类分别位于深圳市、广州市、佛山市周围。结合广东省的人均GDP，深圳、广州及佛山的人均GDP分别为广东省的第一、二、三名，数值上如图表所示

|  |  |  |  |
| --- | --- | --- | --- |
| 城市 | 深圳 | 广州 | 佛山 |
| 人均GDP 元/年 | 5785.6 | 5288.8 | 2153.98 |
|

根据图表结合未完成订单的分布，我们大致确定了一个城市中的未完成概率大致与

其经济繁华程度相关联，在聚类图中，类2为深圳市，类3为广州市，类1位佛山市，据此确定了订单的未完成概率与订单所在城市的繁华程度呈现一定的比例关系。

（二）价格因素

对于拍照赚钱任务，价格是一个决定性因素，一个会员在接到当前单后，在执行拍照赚钱的任务过程会有一个决策，会员会衡量当前的价格是否能够给他带来满意的收益，可以大致确定价格和会员完成订单的概率呈正比关系。数值关系大致如下



其中指的是当前价格下完成任务单的概率， 表示当前价格比例系数，

表示的是完成当前订单可以获得的收益。

（三）距离因素

距离对于订单完成与否的影响规律与价格的相反，价格高人会更愿意去完成，距离远人会考虑是否有必要去完成当前的订单，二者呈现反比关系。数值关系大致如下



其中指的是当前距离下完成任务单的概率 表示当前距离比例系数， 表示的是会员到当前订单的距离。

### 5.2.2 模型建立

(一) 蒙特卡洛法模拟每个会员选单过程

会员选单成功

对于每个会员，我们采用蒙特卡洛模拟法来进行，题目中说会员的限额影响任务的分配，结合此点以及第一问中拟合所得到的数据，我们根据当前选择轮次本会员占附近5km内所有会员限额的比例作为本会员能够选择到当前订单的概率，数学表达式为



表示第i个会员选择任务成功的概率，表示当前会员的限额， 表示当前会员5km范围内相邻的所有会员的限额和。

（二）蒙特卡洛法模拟每个会员订单的完成情况

会员完成概率

根据对附件一中未完成订单分布的可视化以及对其的k-means聚类，我们确定了影响一个任务能否完成的因素有以下三个：

1. 地理位置
2. 任务价格
3. 会员与任务之间的距离

我们假设这三个因素对于任务能否完成的影响是相互独立的，假设会员i完成订单j的地理位置影响的概率为Plij，任务价格的影响概率为Ppij，会员与任务之间距离影响的概率为Pdij，则会员i完成任务j 的概率为P ij = Plij ·Pdij ·Ppij

1、地理位置

对于地理位置对任务是否完成概率的影响，我们已经得到了位于深圳的任务完成的概率低，广州较低，佛山稍低，我们获取了附件一中未完成订单中在这三座城市中聚集的经纬度范围，在后面的判断中如果订单位于这三个城市中，分别取得如下概率



其中Plij为会员i位置如上述4中的概率

2、据任务距离

对于距离对任务是否完成概率的影响，我们求出来所有会员到所有任务的距离，经过统计距离的分布如下表所示

根据拍照赚钱的真实app，我们可以得到会员只能够选择距离自己30km内的任务，app内图片如下图所示



图 app内距离限制

根据确定的距离对于完成的影响是负向的，我们大概确定距离对于任务能否完成符合如下的函数关系式



数值上面的体现即为下面的分段函数



其中为会员i完成距离的任务j的概率

3、任务价格

对于价格对于任务完成与否的影响，因为我们已经结合实际app对于距离的限制将距离限定在30km内，根据平时的生活常识，价格对于任务能否完成的影响应比距离要显著，且这样能更方便的比较新的定价对于所有任务完成率所带来的影响，根据前面分析得到的价格和任务完成率的正比关系，我们直接给出其分段函数如下



其中表示会员i完成价格为的任务j的概率

### 5.2.3 模型求解

具体的模拟过程，在选单过程，我们使用了matlab中自带的rand函数生成一个(0,1)之间的随机数t，此人的选单成功的概率Psi作比较如果t < Psi，就选取此订单，因为生成的随机数是服从均匀分布X~U(0,1)，概率正好与其相同。每个会员的完成过程，通过上面确定的模型，得到了Plij、Pdij、Ppij，完成的概率为P =Plij·Pdij·Ppij，同样使用matlab生成一个服从均匀分布的随机数t，通过大小的比较得到当前的任务是否能够完成，为了避免模型的偶然性，我们对于原来的方案和新的定价方案都模拟1000次，取其平均值作为作为模拟得到的完成订单数。会员选单的模拟系统流程图大致如下

会员选单模拟开始

dij <= 30

根据会员开始批次的大小分批模拟每个会员

当前批次的会员按照信誉值从大到小进行

遍历所有订单判断到此距离dij

False

此订单是否被选

False

True

True

将当前会员能够接的订单按照价格第一关键字、距离第二关键字从大到小排序

遍历有序的订单序列，得到三个影响因素的概率

将当前订单加入到当前会员待完成容器标记当前订单已被选择

当前订单概率Psi与

均匀分布随机数t比较,

Psi < t ?

False

根据当前订单的经纬度范围确定其地理位置影响概率Plij

根据当前订单与会员的距离确定距离影响概率Pdij

根据当前订单价格确定其价格影响概率Ppij

生成均匀分布随机数t

Plij · Pdij ·Ppij < t

当前订单完成，进行记录

确定最大迭代次数，开始迭代

TRUE

当前批次是否结束

TRUE

订单是否已全部被标记

FALSE

True

累计完成量

FALSE

是否达到最大迭代

FALSE

FALSE

取得平均值

程序结束

### 5.2.4 模型结果

根据原订单的价格，为了避免随机概率的偶然性，进行1000次蒙特卡洛模拟，取这1000次完成任务数量的平均值，得到一次模拟的平均完成任务数量为525.861 ，完成率为62.9773%。与附件一中给出的订单，我们统计得到其完成的订单数量为522，完成率为62.5149%，得到的任务完成数量和完成率几乎一致，基于这样的结果我们认为当前的模拟系统是可行的。

新任务定价方案的完成数664.2010

新任务定价方案的完成率79.55%