摘要：

针对问题一，要求我们在确定的变量阶段数和每个阶段的群组数目大小的前提下，做有关测试总数，假阳性率和假阴性率的决策目标，建立数学模型，并进行初步计算和检验。首先考虑决策目标，包括检测总数、假阳性率、假阴性率这三个方面。其次对于两个变量阶段数和每个阶段的群组大小进行考虑。再次用总体样本的特征估计部分样本，用香港地区的确诊率来模拟样本中的患病人数。患病人数随机分布在不同数量的小组决定了检测总数，因此我们利用期望表示平均检测次数。最后我们对三个决策目标进行灰色关联综合测评。通过调整阶段数、分组个数来减小综合测评指标，最终得到最优的群组检测方案。其中对于检测总量为20000，患病率为0.001，最佳的分组为：，检测次数为：，综合评分为：。

摘要：

针对问题二，问要求我们针对无症状受感染者进行检测，建立数学模型找到合理的途经减少人群中的无症状感染者的数量。首先要确定减小无症状感染者对人群影响的方法。其次在本文建立基于传染病模型中的SIR模型，该模型依赖于无症状人群、阳性人群、阴性人群之间的转化关系，以无症状人群数量处于较少状态的时间长短作为方案优劣标准。最后制定有关于两种方案的参数变动比例为30%与60%，观察结果图像，我们发现实施方案前后无症状人群不存在时实施方案比原计划所需的时间要少。

假设：

1、假设仪器存在误差是稳定的。

2、假设将样本融合在一起，其性质不会发生改变。

3、假设一旦被确诊为阳性，会被医院收容，不具有传播属性。

4、假设确诊为阴性的个体，不会限制其活动。

问题分析：

为了在相同数量资源的条件下提高使用效率，我们考虑使用群组检测方法。问题一要求我们在确定的变量阶段数和每个阶段的群组数目大小的前提下，做有关测试总数，假阳性率和假阴性率的决策目标，建立数学模型，并进行初步计算和检验。

首先考虑决策目标，为了更加快捷地进行新冠疫情的检测，我们要考虑检测总数、假阳性率、假阴性率这三个方面的因素。关于检测总数，检测总数越少可以节约资源，在资源有限的前提下做到更多人数的检测。关于假阳性率，也就是阴性被判为阳性的百分比，由于分组样本中出现阳性，则认定改组全部成员为阳性，并进行进一步的检测，直到确定出阳性个体为止，假阳性率越低可以减少隔离人数，做到资源的高效利用。关于假阴性率，也就是阳性被判为阴性的百分比，也就是关于实验仪器的误差，减少检测次数可以控制实验出现误差的情况。

其次，对于两个变量阶段数和每个阶段的群组大小进行考虑。对于阶段数要结合实际进行考虑，查找有关资料得出每日最多接受多少阶段的检测，也就是说在一个工作日内可以进行有效检测的循环的次数，该数据作为阶段数变量的取值范围。对于每个阶段的群组大小来说，要对总体样本进行分组检测，由于每个样本是一个整体，具有不可拆分性，为了方便计算我们考虑用检测样本的因子作为分组的数目。

再次，利用统计概率样本检测相关知识，用总体样本的特征估计部分样本，我们用香港地区的确诊率来估计样本中的患病人数。由于患病人数会随机分配在各个群组样本中，患病人数随机分布在不同数量的小组决定了检测总数，因此我们利用期望表示平均检测次。利用随机取样原理，通过多次实验，使其测试总数趋于稳定。

最后我们对三个决策目标进行综合测评，由于三个决策目标之间有联系，利用灰色关联得出综合指标。通过调整阶段数、分组个数和组内样本数量来减小综合测评指标，最终得到最优的群组检测方案。

模型的建立：

***Step1*关于两个变量**

* 阶段数（step）

根据查找的资料显示利用核算检测方法检测一个样本的时长为6小时，一家医院的正常工作时长为9小时。假设医院在疫情期间轮班制，处于合理考虑我们认为一天内可以进行的检测循环为两次即阶段数的取值范围为1-2。由于阶段数为1代表将所有样本都进行一次检验，因此我们主要考虑阶段数为2的情况。

* 每阶段群组大小（N）：

每个阶段的群组大小是总体样本数量的多少。由于要对总体样本进行分组检测，由于每个样本是一个整体，具有不可拆分性，为了方便计算我们考虑分组个数为总体样本的因子。

其中表示第阶段的分组个数，表示第阶段组内样本数量, 第阶段群组数量大小的计算公式为：

***Step2*三个决策目标**

* 检测次数

检测次数的计算步骤如下所示：

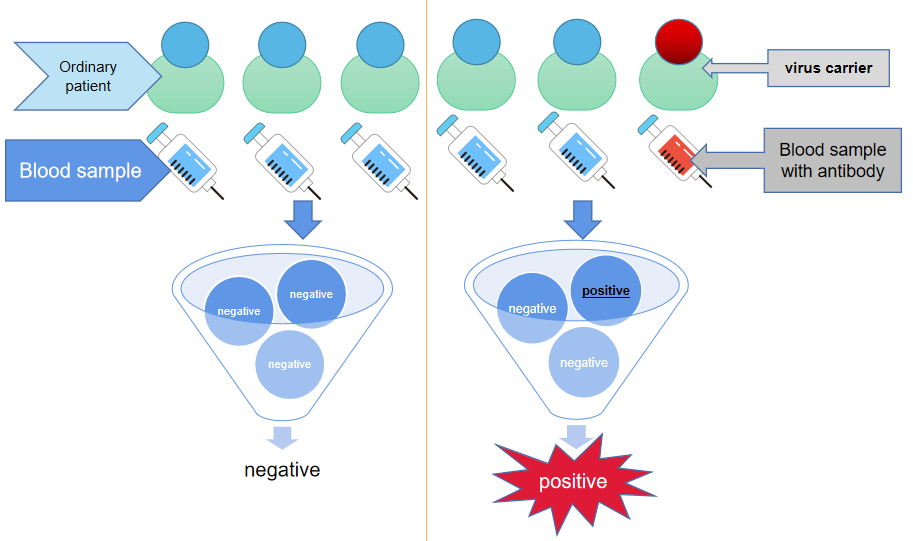
***Step1*确定样本总量的感染人数**

由于样本中感染人数不同，会导致感染人数被随机分配在不同小组中的概率不同，从而得到不同的测试总数。我们设感染率为，则感染人数：

。

***Step2*考虑阶段数为2的测试总数**

考虑阶段数为2情况，对于第一阶段来说，先将第一阶段的样本总数进行分组，将其分为a组，这里的a为N的因子。因此在第一阶段检测的次数为a次。由于感染人数随机的分配在不同组内，对于每个小组来说，只要该组中有一个感染样本就表示该组为感染，随机分配事例图如下所示：



对于分组后小组中有感染样本的小组数目进行讨论。如果感染人数b大于分组个数a，则最多有a个组出现阳性的情况；如果感染人数b小于分组个数a，则最多有b个组出现阳性的情况。

设a组中有c组中有感染人数，检测总数的计算公式如下所示:

由于c的不同导致检测次数不同，感染人数会随机分配在不同数量的组内。由题意可得，一个小组内有感染人数就表明该组全部成员全部感染，进行下一步检测，直到检测出单个样本是否感染，因此我们只需要考虑感染小组个数，因此我们参考假货概率的原理。假货概率原理是说在一堆硬币中有一部分假硬币，其中假硬币的质量比真硬币的质量小，将该堆硬币随机分组，其中存在假硬币的小组质量小于全部真硬币的小组，得到存在假硬币小组的个数概率。利用该概率求出检测次数的期望表示出在不同分组的情况下平均检测次数的大小，期望的公式如下所示：

找出期望最小的分组方式作为阶段为2的最优的群组检测方案。

* 假阳性率

关于假阳性率，也就是阴性被判为阳性的百分比，由于分组样本中出现阳性，则认定改组全部成员为阳性，并进行进一步的检测，直到确定出阳性个体为止，假阳性率越低可以减少隔离人数。设假阳性率为，计算公式为：

* 假阴性率

关于假阴性率，也就是阳性被判为阴性的百分比，也就是实验仪器的误差，设实验仪器出现误差的概率为，假阴性率的计算公式为：

***Step3*综合评价模型**

从决策目标的角度出发，我们选择了测试总数、假阳性率、假阴性率作为评价群组检测方法的优劣的指标。由于各个指标之间相互联系但是有反应不同方面的性质，因此要建立一个科学合理的综合评价模型，我们选用灰色关联模型，具体模型如下所示：

确定总体样本数量，不同的分组会得到不同的检测总数、假阳性率、假阴性率，如下表所示：

|  |  |  |  |
| --- | --- | --- | --- |
| 分组个数 | 检测总数 | 假阳性率 | 假阴性率 |
|  |  |  |  |
| ... | ... | ... | ... |
|  |  |  |  |

* **选取母指标**

在step2中我们求出关于检测总数、假阳性率、假阴性率的公式，我们可以发现假阴性率和假阳性率与检测总数有联系，且检测总数最直观的反映出群组检测的优劣，因此选取检测总数母指标。

* **计算关联系数**

定义母指标与在第点的关联系数为：

其中的计算公式如下所示：

的计算公式如下所示：

的取值范围为,这里我们取。n

* **计算关联度**
* **构建综合评价模型**

从上一步中求出的关联度，计算各个指标的权重：

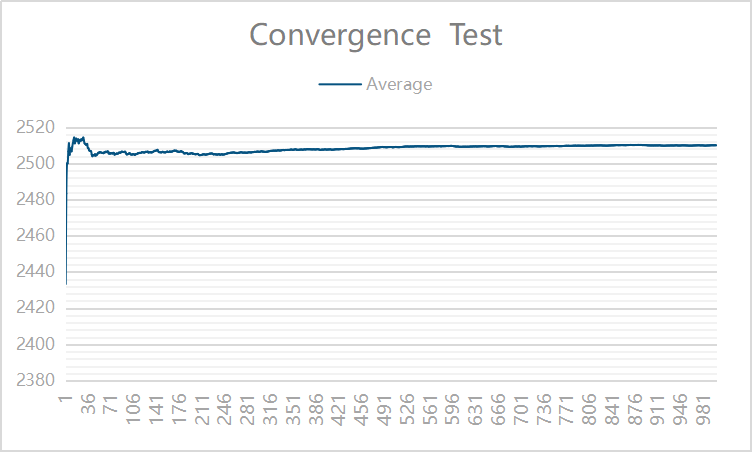
作为各个指标的权重，得到综合指标.

最后综合指数最小所对应的分组数即为总体样本数N最优的分组方法。

模型的求解：

关于感染率的取值时我们用总体样本的特征来估计部分样本，也就是说用香港地区的确诊率来估计样本中的患病人数。根据国家统计局数据显示目前香港感染人数为70000，其中香港总人口为734.567万人。感染率为千分之一，因此我们取=0.001。

计算检测次数时，由于分组后小组出现阳性概率易于计算，由于小组中只要存在阳性个体则表示该小组为阳性，计算存在多少小组为阳性样本的概率过于复杂，但其一定存在，因此我们选用大量随机模拟得到概率和期望，下面我们展示随机次数与期望的图像：



由图像可得，在随机次数为350次数时，检测次数的期望基本趋于一致，在极小的范围内波动。为了减小计算的误差，在计算检测次数的期望时我们取随机次数为10000。

利用模型中假阳性率和假阴性率公式计算出不同分组下的数值。下面举一个样本总数为40000的例子，给出不同分组情况下的数据值如下表所示：

|  |  |  |  |
| --- | --- | --- | --- |
| 分组数a | 检测次数 | 假阳性率 | 假阴性率 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

根据灰色关联模型求出三个指标与母指标之间关联度如下表所示：

|  |  |  |  |
| --- | --- | --- | --- |
| 指标类型 | 检测次数 | 假阳性率 | 假阴性率 |
| 关联度 |  |  |  |

每个指标的权重如下表所示：

|  |  |  |  |
| --- | --- | --- | --- |
| 指标类型 | 检测次数 | 假阳性率 | 假阴性率 |
| 权重 |  |  |  |

得到综合测评指标为：

综合测评指标值越小表示分组方案更加高效。

计算出各项分组的综合指标如下表所示：

|  |  |  |
| --- | --- | --- |
| 分组数 | 综合指标 | 评分名次 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

得到总体样本数为40000的最优分组为400组。

重复上述步骤，得出在不同样本数的情况下最优的分组方案，结果如下表所示：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 检测总量 | 患病率 | 患病人数 | 分组数 | 检测次数 | fake positive | fake negative | score |
| 4000 | 0.001 | 4 |  |  |  |  |  |
| 6000 |  |  |  |  |  |  |  |
| 8000 |  |  |  |  |  |  |  |
| 10000 |  | 10 |  |  |  |  |  |
| 20000 | 0.001 | 20 | 1020 | 1406.0784 | 0.018304 | 0.001712 | 0.97 |
| 30000 |  |  |  |  |  |  |  |
| 40000 | 0.001 | 40 |  |  |  |  |  |

灵敏度分析

在上述模型中，我们用香港地区的确诊率作为检测样本中的患病率，但是对于不同区域其患病率会发生变化。在疫情高发区，患病率会有所增加；在安全区，患病率相对较小。为了应对不同区域的检测，我们将感染率进行调整，得出不同感染率下对应的最优分组策略。下面我们做了总体样本为4000，10000和20000的灵敏度分析

下表为总体检测人数为4000，不同感染率所对应的分组策略：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 检测总量 | 患病率 | 患病人数 | 分组数 | 检测次数 | fake positive | fake negative | score |
| 4000 | 0.001 | 4 |  |  |  |  |  |
| 4000 | 0.005 | 20 | 445 | 619.561 | 0.03864 | 0.00387 | 0.945 |
| 4000 | 0.01 | 40 | 610 | 862.393 | 0.05309 | 0.00559 | 0.929 |

总体检测人数为10000，结果如下所示：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 检测总量 | 患病率 | 患病人数 | 分组数 | 检测次数 | fake positive | fake negative | score |
| 10000 | 0.001 | 10 |  |  |  |  |  |
| 10000 | 0.002 | 20 | 755 | 1014.8675 | 0.023987 | 0.002305 | 0.96229324 |
| 10000 | 0.005 | 30 | 1175 | 1589.7234 | 0.036472 | 0.003679 | 0.94502669 |
| 10000 | 0.01 | 100 | 1529 | 2159.6736 | 0.053067 | 0.005594 | 0.92893737 |

总体检测次数为20000，结果如下所示：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 检测总量 | 患病率 | 患病人数 | 分组数 | 检测次数 | fake positive | fake negative | score |
| 20000 | 1020 | 1406.07 | 0.01830 | 0.00171 | 0.97213 | 1020 | 1406.07 |
| 20000 | 1435 | 1982.87 | 0.02539 | 0.00243 | 0.96232 | 1435 | 1982.87 |
| 20000 | 2287 | 3138.77 | 0.0375 | 0.00377 | 0.94520 | 2287 | 3138.77 |
| 20000 | 3179 | 4394.47 | 0.05077 | 0.00539 | 0.92900 | 3179 | 4394.47 |

问题分析：

问题二要求我们针对无症状受感染者进行检测，建立数学模型找到合理的途经减少人群中的无症状感染者的数量。

首先要确定减小无症状感染者对人群影响的方法，通过查阅文献，得到减小无症状感染者对人群影响的途经有以下两种，一是进行大面积的核酸检测，查找出无症状人群中携带新型冠状病毒的个体；另一种是对没有进行检查的居民实行限制出行的政策。

其次为了综合以上两种方案，本文考虑将全体人群分为无症状人群，确证携带病毒的人群和确证不携带病毒的健康人群，基于传染病模型中的SIR模型建立关于应对无症状感染者处理模型，该模型依赖于三个人群之间的转化关系。改变转化关系中的参数来模拟在两种方案对无症状人群的控制和处理，以无症状人群数量处于较少状态的时间长短作为方案优劣标准。

最后制定有关于两种方案的参数变动比例为30%与60%，参数的变动比例越大，方案的影响程度越广，因此比较不同方案强度实施前后的效果，围绕该效果说明我们应该如何应对无症状感染者继续传播疫情。

**5.2问题三模型**

通过查阅文献，得到减小无症状感染者对人群影响的途经有以下两种：

1.进行大面积的核酸检测，查找出无症状人群中携带新型冠状病毒的个体。

2.对没有进行检查的居民实行限制出行的政策。

**5.2.1基于SIR的无症状传播模型**

在建立该模型前，我们先确定人群的类型以及无症状人群接受核酸检测的概率，具体讨论如下所示：

* 人群类型

在疫情的传播中，不同类型的人群对疫情的传播有不同类型的影响。查阅资料，我们发现该模型可以类比SIR模型，在SIR模型中将人群分为易感者、感染者、治愈者。类似地，我们将人群分为S、I、R三部分，具体含义如下所示：

|  |  |  |
| --- | --- | --- |
| 符号 | 含义 | 简称 |
| S | 无症状人群（有感染者和健康者） | 无症状人群 |
| I | 确诊为阳性的人群 | 阳性人群 |
| R | 确诊为阴性的人群 | 阴性人群 |

设t时刻无症状人群占总人群比例为s(t), 阳性人群占总人群比例为i(t)，阴性人群占总人群比例为r(t)，记初始值为，,。由于总人群数不变，三种人群的比例显然有如下公式：

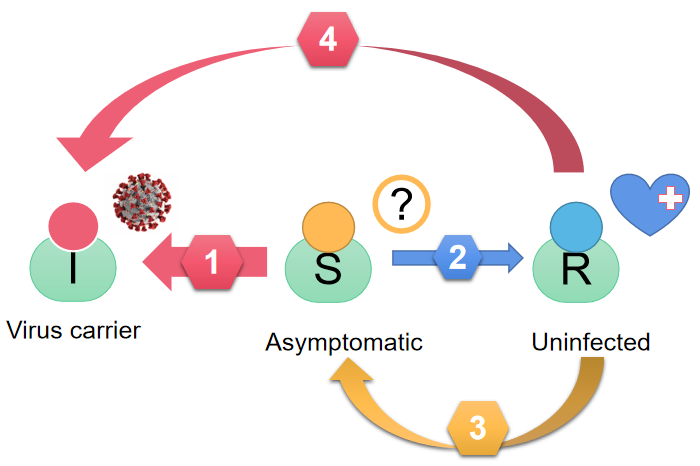
* 无症状人群接受核酸检测的概率

在政府进行大面积的核酸检测时，每个无症状人群接受检测的概率为：

其中A表示每日可检测的样本数，M表示总体人数。换句话说，上式概率为可检测的样本数占总无症状人群的占比。如果扩大检测量，表示A增大，每个无症状样本接受检测的概率就越高。

**5.2.2 无症状传播模型的建立**

首先考虑无症状人群S、阳性人群I和阴性人群R之间的关系，通过分析得到三者之间的转化关系，转化关系示意图如下所示：



无症状人群S检测后诊断为阳性转变为阳性人群I。

无症状人群S检测后诊断为阴性转变为阴性人群R。

阴性人群I容易被无感染人群S中携带病毒的人传染，但是并没有显示症状变为无症状人群S。

阴性人群I容易被无感染人群S中携带病毒的人传染，显示症状后及时就医转变为阳性人群S。

其次对四种转化关系进行分析，得出微分方程，具体步骤如下所示：

考虑情况，阳性人群的增加来源于无症状人群S按照概率进行检测并且结果为阳性和阴性人群I被感染出现症状转变为阳性人群。设检测为阳性的概率为,被感染出现症状的概率为综合以上信息，得到的变化率为：

考虑情况，对于无症状人群来说，无症状人群的人会减少分别流入阳性人群和阴性人群。设检测为阴性的概率为，则得到的变化率为：

由,简化后的变化率为：

考虑情况，对于阴性人群来说，人数的增加来源于无症状人群检测后为阴性的人数。人数的减少为两个方面，一个是被传染出现症状及时就医变为阳性人群，另一个是被传染没有出现症状转变为无症状人群。同理可得，的变化率为：

综合以上讨论，得到基于SIR的无症状传播模型

**5.2.3 无症状传播模型的求解**

利用matlab对该模型进行求解，我们通过文献资料对模型中的参数进行赋值，赋值结果如下所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 参数 | 数值 | 参数 | 数值 | 参数 | 数值 |
|  | 0.6 | M | 7000000 |  | 0.99 |
|  | 0.1 | A | 10000 |  | 0.63 |
|  | 0.3 |  | 0.01 |  | 0.03 |

上述两个干预方法是对参数A和参数进行调整。关于参数A的分析，由于方案要大面积进行核酸检测，因此每日检测人数要有所上升。关于参数的分析，方案要求隔离未进行检测的人群，为了减小无症状人群中携带病毒的人将病毒传染给阴性人群。对两个参数进行调整，调整的比例如下所示：

|  |  |  |
| --- | --- | --- |
| 参数 | 变化情况 | 变化情况 |
| A | 增加30% | 增加60% |
|  | 减少30% | 减少60% |

得到的图线如下所示：

我们发现，变化后的图线比没有变化的图线，无症状人群减小到零所需要的时间更少。同时变化比例越大所需时间越少。

综上所示，我们发现大面积的核酸检测和控制未检测人员的外出可以应对无感染人群中携带病毒的个体对整个人群进行传播。