

基于数学期望和排队论的核酸混检方案的设立

摘要

核酸检测是一种迅速切断病毒传播途径的方式，常规的检测办法是逐人依次检测，对样本量比较小的情况，这种方法可以保证检测结果及时有效，而对量大且患病率很低的疾病样本进行检测时，单样本检测耗时较多，影响检测速度，不利于迅速切断病毒传播。针对大规模人群进行全面核酸检测时，混合采样检测的方法可以在巨大的工作量下节省时间。

针对问题一，基于离散型变量的数学期望模型和排队论模型建立核酸混合采样检测的数学模型。首先考虑大规模进行核酸检测时要追求高效，因此设置每组每人检测次数为变量 x ，建立其数学期望模型。然后考虑到在大规模核酸检测过程中存在的排队问题以及新冠病毒的高传播能力，建立排队论模型，并以系统中平均等待时间为参考指标。最后综合进行考虑，以每组每人平均检测次数和平均等待时间为首要参考指标，结合上述两个模型设立最优约束条件进行求解。

针对问题二，考虑到深圳市与上海市在本轮疫情中的不同表现，本研究将选择这两座城市中发生的真实案例来代表不同的疫情严重程度，以完成两个案例的核酸混检方案为目标，考虑每组混检人数的选择与疾病的患病率之间的关系，以及病毒的传播能力，制定科学的混合采样检测方案。以国家卫健委发布的《关于印发区域新型冠状病毒核酸检测组织实施指南（第三版）的通知》为参考，规定相关参数的取值范围。

针对问题三，考虑到奥密克戎病毒，潜伏期长，传播能力的特点以及核酸检测过程中会出现的假阴性问题，人工操作等其他误差，我们基于问题一所提出的数学期望模型进行了拓展，以患病率和错误率作为两个参数，通过非线性规划方法确立我们的混检方案。

目录

1.问题的重述	1
1.1 背景分析	1
1.2 问题的提出	1
2.问题的分析	2
2.1 问题一的分析	2
2.2 问题二的分析	2
2.3 问题三的分析	3
3.模型的假设	3
4.符号说明	3
5.模型的建立与求解	4
5.1 问题一模型的建立与求解	4
5.1.1 离散型随机变量的数学期望模型的建立	5
5.1.2 排队论模型的建立	7
5.1.3 核酸混检模型的建立	9
5.2 问题二模型的建立与求解	9
5.2.1 数据与资料收集	9
5.2.2 核酸混检方案的解决	11
5.3 问题三模型的建立与求解	16
5.3.1 模型的假设	16
5.3.2 期望模型的建立	16
6.模型的分析与评价	19
6.1 模型的分析与评价	19
6.1.1 模型的优点	20
6.1.2 模型的缺点	20
6.2 模型的推广	20
7.小结	20
8.参考文献	20

1. 问题重述

1.1 背景分析

2019 年底，新冠疫情开始席卷全球，由于新型冠状病毒极强的变异性，直至如今，我们的生活仍受新冠疫情的困扰。奥密克戎（Omicron）作为新型冠状病毒（COVID-19）的变异种之一，其传播能力、传播速度大幅提升，平均传播速度是一代 2.7 天，三代就接近 9 天。2022 年初，上海疫情再次爆发，根据上海市疾控中心病毒基因测序结果显示，导致本轮上海疫情的新冠病毒，绝大部分为奥密克戎 BA.2 型和 BA.2.2 型，这一波疫情的发生、发展，也完全符合人类现已掌握的奥密克戎变异株的特点，这也是和国内外其他地区由奥密克戎变异株所导致的疫情特点基本上是一致的。

核酸检测作为一种迅速切断病毒传播途径的方式，是通过实验室方法进行检测、筛查，发现新冠肺炎确诊病例、疑似病例和无症状感染者，确保“早发现、早报告、早隔离、早治疗”。在医学疾病筛查检测中，常规的检测办法是逐人依次检测，也称作单样本检测。对样本量比较小的情况，单样本检测可以保证检测结果及时有效，而对量大（如百万级以上样本），且患病率很低的疾病样本进行检测时，采用单样本检测进行筛查不仅需要大量的人力、物力和财力，而且比较费时，同时也可能因为耗时较多，影响检测速度，不利于相关单位及时制定措施。因此，针对大规模人群进行全面核酸检测时，如何在巨大的工作量下节省时间，为控制病毒传播争分夺秒成为了一个难题。混合采样检测的方法，即将若干人组成一组，把他们的采样样本混合在一起进行检测，能大幅降低总体检测时间，有效控制新冠疫情的传播。

1.2 问题的提出

根据混合采样检测的性质可知，若一组样本的检测结果为阴性，则这组样本对应的被检测者都是阴性的；反之，若一组样本的检测结果为阳性，则这组样本中至少有 1 人是阳性的，那么就需要对这组样本的所有被检测者重新进行检测。考虑到病毒极强的传播能力和传播速度，在重新检测时，将采用单人单管检测的方法来确定这组样本中的所有阳性被检测者。因此，确定一个合适的混检方案能有效地提高核酸检测效率、控制病毒传播。

本题就如何确定合适的混检方案提出以下问题：

问题一：建立核酸混合采样检测的数学模型，并根据实际情况，以核酸混合检测平均队长、平均排队时间以及每组每人的平均检测次数为首要参考点，说明较为理想的进行混检的每组人数。

问题二：通过收集患病率、核酸检测点安排、核酸检测时间、检测效率等相关数据，根据问题一所建立的模型，考虑具体步骤、混检人数、排队时间等方面，给出合适的混检方案。

问题三：根据实际情况，考虑到现实生活中核酸检测的各个环节难免出现差错导致假阳性、假阴性等情况，出现疫情时往往需要进行多轮核酸检测。通过进行多目标规划来建立多轮检测的数学模型，分析说明在进行多轮检测时是否需要调整第二问所确定的混检方案，若需调整，给出调整后的混检方案。

2、问题分析

混合采样检测是一种能够提高检测效率的检测方法，虽然当每组混检人数 k 越大时，第一轮检测只需要检测较小数量的组数即可得到结果，大大节省了核酸检测的时间。但是每组混检人数过大时，一旦某一组出现问题，就需要对这组里的 k 各人进行单人单采，相当于该组一共进行了 $k+1$ 次核酸检测。并且考虑到阳性患者是随机分布在各个组里的，需要重新进行单采核酸检测的组数是随机的，因此并非每组混检人数 k 越大越好。

2.1 问题一的分析

问题一要求建立核酸混合采样检测的数学模型并找出较为理想的每组核酸混检人数，考虑到题目要求仅为求得较理想的方案，为简化模型，在第一问时将不考虑存在“假阴性”的情况，即假设在这一问核酸检测过程中不存在任何差错，检测结果 100% 准确。

结合实际情况与大众心理分析，以总体核酸检测人数、总体检测时间以及平均排队时间为指标，寻找较为理想的核酸混合采样方案。建立排队论模型，通过变换每日核酸检测时间，以及服务台的数量，以检测点预期人数和计划检测点数为指标，选得合适的每日核酸时间及服务台数量，对二者进行相关的约束，以此得出系统平均等待队长，顾客平均等待时间。同时建立离散型随机变量数学期望模型，对两个模型同时进行目标优化，求解出理想的每组混检人数。

2.2 问题二的分析

在 2022 年的这轮疫情中，由于香港疫情的爆发，深圳、上海等城市肩负着支援香港、承接香港航班的任务，展现出一线城市的担当。深圳在“外防输入”、“内防扩散”的压力下，凭借合理、有效的抗疫政策成为了中国科学抗疫模式的杰出代表，但上海的疫情却迟迟没有得到有效的控制。因此，本题将以深圳和上海的疫情数据为代表，对不同疫情严重情况的核酸混检方案进行合理的假设估计。

根据国家卫健委发布的《关于印发区域新型冠状病毒核酸检测组织实施指南（第三版）的通知》等相关政府文件，全国以及深圳市在此轮疫情中的实际防疫政策、累计感染人数、常住总人口等信息，对问题一中设置的相关参数做出合理假设估值，并控制变量进行比较，筛选出合适的检测点数量，并以此与第一问中离散型随机变量的数学期望结合，通过计算得出最优的核酸混合采样检测配置方案。

2.3 问题三的分析

核酸检测的准确率也受到如标本取材时采集到的病毒数量、样本存放的温度和时间、检测过程的规范性等有关因素的影响。并且，考虑到现实生活中流行的奥密克戎病毒具有极强的传播能力，未感染的人也有被感染的风险，这就会导致每人每天的身体状况可能会有所不同。这个时候就需要进行多轮检测，以确保能彻底地检测出人群中存在的所有阳性感染者，完全控制住新冠病毒的传播。

在多轮检测中，考虑到不同地区的疫情严重程度不同，针对不同的疫情严重程度等级，考虑患病率和核酸检测的错误率，在问题一设立的数学模型的基础上

进行拓展，通过建立多元变量方程，以非线性规划的方式设置不同的混检方案以形成最为合理的核酸检测方案。

3、模型假设

基于以上分析，我们提出以下合理假设：

1. 根据国家药品监督管理局医疗器械技术评审中心发布的《2019 新型冠状病毒核酸检测试剂注册技术审评要点》，假设核酸检测准确率是 100%，即不考虑“假阳性”、“假阴性”的情况；
2. 假设新冠疫情累计感染人数的统计结果数据能真实反映该地区的患病率，数据来源可靠；假设感染新冠疫情的人群占总体人群的比例为 p ，则新冠疫情的患病率为 p ；
3. 假设在排队进行核酸检测时，有被感染的风险，且其余时候无感染风险；
4. 假设传染病空气动力学模型忽略了潜伏期患者导致的传播风险；
5. 假设本模型尚未考虑无症状感染者的影响，未来可以在本模型的基础上考虑无症状感染者，使得模型更加贴合实际；
6. 假设疫情传播过程中病毒未发生显性突变。

4、符号说明

根据本文需要，我们引入如下符号说明：

表 1 模型中符号的说明

符号	说明
K	每组混样检测人数
p	新冠病毒患病率
X	每组中每个人需要的检测次数
c	服务台数量
ρ	服务强度
λ	顾客到达率
μ	平均服务率
T	日工作时间
s	核酸检测点数量
q	检测点平均日容量
R	所需医护人员数量
N	需要进行核酸检测的总人口数
r	核酸检测错误率

5、模型的建立与求解

5.1 问题一模型的建立与求解

所谓混合采样检测就是把要检测的样本混合在一起做一次检测，是一种常用的提高检测效率的检测方法。面对大量的样本，在具体实施时通常会将检测样本分成若干组，对每组的 k 个样本进行混合采样检测，根据检测结果判断这组样本的检测结果，即用混合采样检测结果代替混样个体样本检测结果。如果混合采样检测结果呈阳性，就说明这组样本中至少有一样本感染了新冠病毒，此时，需对这 k 个样本进行单人单管检测来确定该组样本每个个体的结果。这样，对该组样本实际上做了 $k + 1$ 次检测。因此，如果 k 取值不当，不仅不能提高核酸检测的效率，反而会加大核酸检测的工作量，影响疫情防控。因此，可以使用核酸检测次数为指标来衡量 k 的取值。设 X 为每组中每个人需要的检测次数，易知 X 是离散型随机变量，则 X 的数学期望 $E(X)$ 可以表示每组中每个人的平均检测次数。

事实上，在实际检测中， k 的选择不但与疾病的患病率 p 有关，而且还与检测的准确性等因素有关，需要综合考虑，才能制定科学的混合采样检测方案。但由于核酸检测非常精准，出现“假阴性”的概率较低，且一旦出现问题难以检测出，故本题在建立模型时将不考虑“假阴性”的情况，下面进行具体分析：

假设某种疾病在人群中的感染率为 0.1% ，医院现有技术对于该疾病检测的准确率为 98% ，即在已知某人受到感染的情况下，该人的检查结果有 98% 的可能性为阳性，正常未被感染的人的检测结果是阳性的概率为 1% ，即正常人有 99% 的可能性被检查为正常。

这时，如果要计算感染某人是否被感染的概率就需要考虑到先验概率和后验概率这两个问题。先验概率即为“已知某人感染疾病，该人检测结果为阳性的概率”；后验概率则为“已知某人检测为阳性的条件下，该人被感染的概率”。

假设用 A 表示某人感染该疾病，用 B 表示医院检测的结果为阳性，先验概率为 $P(A) = 0.1\%$ ，而现随机抽检一人，医院给出的检测结果为阳性，则此人被感染的概率即为后验概率，即 $P(A|B)$ 。根据贝叶斯公式求得此人检测结果为阳性，并且感染的概率为：

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

其中： A^c 为 A 的对立事件， $P(A^c) = 1 - P(A)$ ； $P(A)$ 表示被检测者感染的概率； $P(A^c)$ 表示被检测者未感染的概率； $P(B|A)$ 表示已知感染的情况下检测为阳性的概率； $P(B|A^c)$ 表示已知未感染的情况下检测为阳性的概率。

从贝叶斯定理的角度来看，随机抽检一个被测者，由于检测之前此人的信息并不充分，故未检测之前有真阳性 $P(A|B)$ 、假阳性 $P(A^c|B)$ 、真阴性 $P(A^c|B^c)$ 和假阴性 $P(A|B^c)$ 四种可能性，这些可能性的大小由该疾病的感染率及检测技术的准确率决定。其概率见下表

表 2 检测结果概率

	感染	未感染
阳性	0.00098（真阳性）	0.00999（假阳性）
阴性	0.00002（假阴性）	0.98901（真阴性）

由表可知，随机抽检一个被测者，其检测结果为阴性包含两种可能性：真阴性和假阴性。即使检测结果为阴性，也不能代表被测者未被感染，这就是所谓的“假阴性”，二者发生的概率分别为 0.98901 和 0.00002，假阴性的概率仅为真阴性概率的十万分之一，说明虽然存在“假阴性”的现象，但它的概率非常小，可以忽略不计，故本题在建立模型时将不考虑“假阴性”的情况。

5.1.1 离散型随机变量的数学期望模型的建立

假设现有一批需要进行核酸检测的样本，每个样本的检测结果呈阳性的概率为 p ，即这批样本中患病率为 p ，样本检测结果为阴性的概率是 $q = 1 - p$ 。将这批样本分成若干组，每组 k 个。

由于核酸检测方式对新型冠状病毒，即要检测项的灵敏性比较高。国家卫健委曾在官方文件中，对“混检”有两种模式的可靠性进行回应，表示“混合采样”不会影响核酸检测的敏感性。

那么，如果混合样本中有一个样本呈阳性，则混合检测结果也呈阳性；如果混合检测结果呈阴性，则认为该组人的检测结果都呈阴性。

若用 X 表示每组中每个人需要的检测次数，则 X 是离散型随机变量，且

$$X = \begin{cases} \frac{1}{k}, & \text{该组混合样本结果为阴性} \\ 1 + \frac{1}{k}, & \text{该组混合样本结果为阳性} \end{cases}$$

由此易得出 X 的分布律为：

$$P\left\{X = \frac{1}{k}\right\} = q^k; \quad P\left\{X = 1 + \frac{1}{k}\right\} = 1 - q^k$$

这样，随机变量 X 的数学期望 $E(X)$ 就能够表示每组中每个人的平均检测次数：

$$E(X) = \frac{1}{k} \cdot q^k + \left(1 + \frac{1}{k}\right)(1 - q^k)$$

化简可得：

$$E(X) = 1 + \frac{1}{k} - q^k$$

为了有效降低检测工作量，对给定的 $q = 1 - p$ ，只要选取每组人数 k ，使得：

$$E(X) = 1 + \frac{1}{k} - q^k < 1$$

根据题目要求，为了得到较为理想的合适的每组混检人数， $E(X)$ 表示的每组中每个人的平均检测次数的取值应该越小越好。

由此，问题就转化为：在排队论中系统平均等待队长的约束下，求 $1 + \frac{1}{k} -$

q^k 的极小值。为了处理方便，可设 $f(x) = 1 + \frac{1}{x} - q^x$ ($x > 0$)。分别选取 $p=0.3, 0.2, 0.1, 0.01, 0.001, 0.0001$ ，探究在不同患病率的情况下，每组混检人数与每个人平均核酸检测次数之间的关系。

用 MATLAB 软件编程分别绘制函数 $f(x)$ 的曲线，并计算得出不同患病率 p 值下最优的每组混检人数 k 以及每组中每个人的平均检测次数 $E(X)$ 。

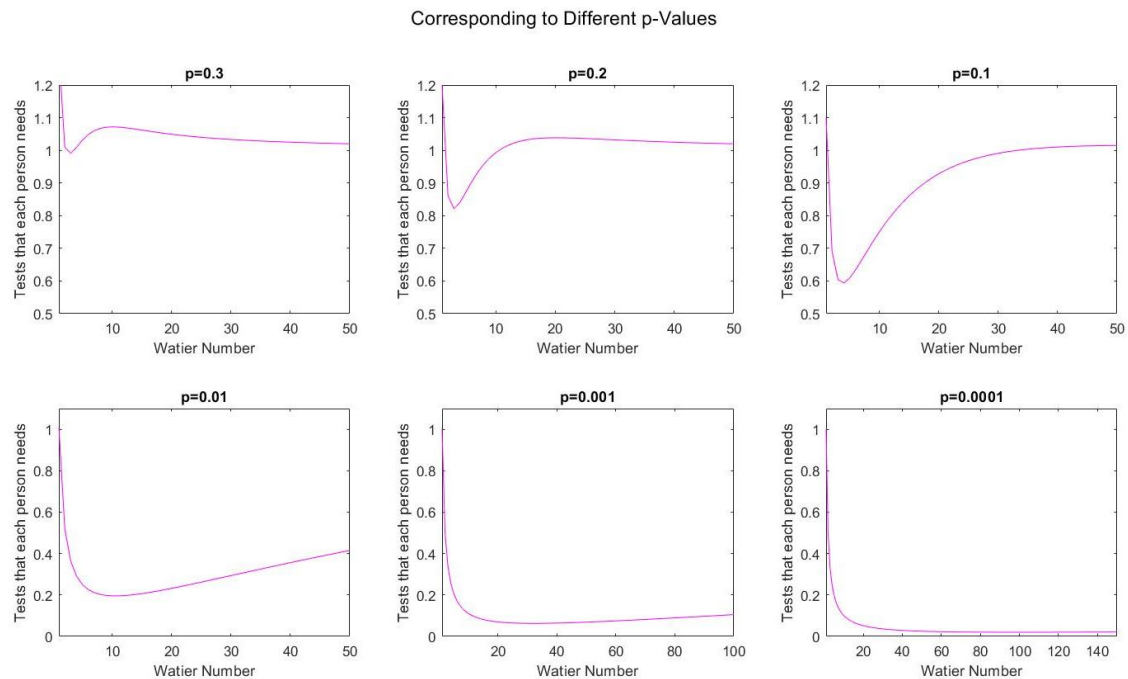


图 1 不同患病率 p 对应的函数 $f(x)$ 的曲线

表 3 不同 p 值对应的最优 k 值以及期望 $E(X)$

p	0.3	0.2	0.1	0.01	0.001	0.0001
k 的最优值	4	4	5	11	32	101
$E(X)$	0.99033	0.82133	0.5939	0.19557	0.06276	0.01995

由图可知，以患病率 p 作为参数时，求 $\min \left\{ f(x) = 1 + \frac{1}{x} - q^x \right\}$ ($x > 0$) 时所对应的 x 即为所需每组混检人数 k 。但是考虑到如果没有合理优化设置的核酸检测点，将会导致大量的人员聚集，从而增加病毒的传染几率，也会使居民因等待时间过长而引发不满，为了解决这一问题我们对此又建立了相关的排队论模型对每组混检人数 k 加以约束。

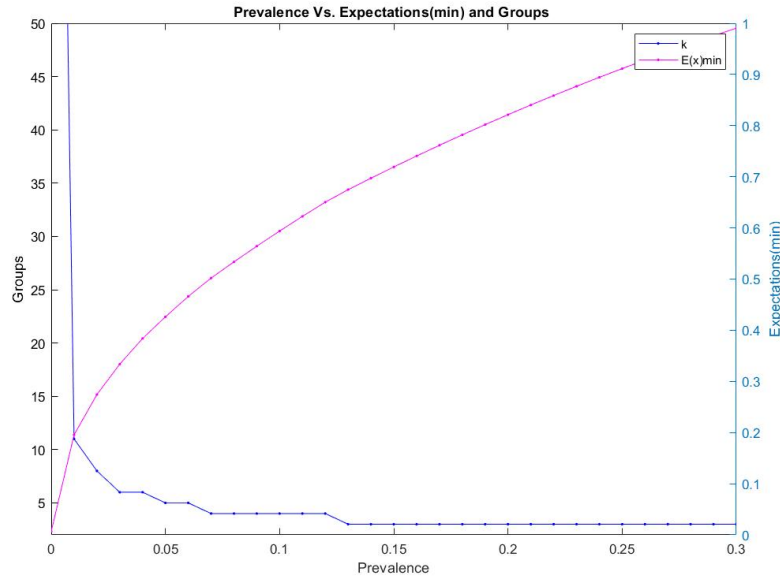


图 2 不同 p 值对应的最优 k 值以及期望 $E(X)$

5. 1. 2 排队论模型的建立

由于顾客到达和服务时间的随机性，可以说排队现象几乎是不可避免的。排队论（Queuing Theory）也称随机服务系统理论，就是为解决上述问题而发展的一门学科。考虑到现实生活中在进行核酸检测时，不论是单人单采还是多人混采都需要经历排队的过程，因此在建立第一问核酸混样检测的模型时，可以把排队论的模型也考虑进其中。

排队论模型中，定义凡要求服务的对象统称为顾客，为顾客服务的人或物称为服务员，由顾客和服务员组成服务系统。各个顾客从顾客源出发，随机地来到服务机构，按一定的排队规则等待服务，直到按一定的服务规则接受完服务后离开排队系统。其中规定，

（1）排队规则（指到达排队系统的顾客按怎样的规则排队等待）：

损失制，即当顾客到达时，所有的服务台均被占用，顾客随即离去；等待制，即当顾客到达时，所有的服务台均被占用，顾客就排队等待直到接受完服务才离去；混合制，介于损失制和等待制之间的是混合制，即既有等待又有损失。有队列长度有限和排队等待时间有限两种情况，在限度以内就排队等待，超过一定限度就离去。

（2）服务机构：

单服务台；多服务台串联，即多个服务台依次为同一顾客服务；多服务台并联，即每个服务台同时为不同顾客服务；混合型。

（3）服务规则（指为顾客服务的次序）：

先到先服务；后到先服务，如情报系统中，最后到的情报信息往往最有价值，因而常被优先处理；随机服务，服务台从等待的顾客中随机地取其一进行服务，而不管到达的先后；优先服务，如医疗系统对病情严重的病人给予优先治疗。

设 $N(t)$ 表示在时间区间 $[t, t + \Delta t]$ 内到达的顾客数 ($t > 0$)，令 $P_n(t_1, t_2)$ 表示在时间区间 $[t_1, t_2]$ ($t_2 > t_1$) 内有 n ($n \geq 0$) 个顾客到达的概率，即

$$P_n(t_1, t_2) = P\{N(t_2) - N(t_1) = n\} (t_2 > t_1, n \geq 0)$$

当 $P_n(t_1, t_2)$ 满足以下三个条件时，

(1) 在不相重叠的时间区间内顾客到达数是相互独立的，我们称这性质为无后效性。

(2) 对充分小的 Δt ，在时间区间 $[t, t + \Delta t]$ 内有一个顾客到达的概率与 t 无关，而约与区间长 Δt 成正比，即

$$P_1(t, t + \Delta t) = \lambda \Delta t + o(\Delta t)$$

其中 $o(\Delta t)$ ，当 $\Delta t \rightarrow 0$ 时，是关于 Δt 的高阶无穷小。 $\lambda > 0$ 是常数，它表示单位时间有一个顾客到达的概率，称为概率强度。

(3) 对于充分小的 Δt ，在时间区间 $[t, t + \Delta t]$ 内有两个或两个以上顾客到达的概率极小，以致可以忽略，即

$$\sum_{n=2}^{\infty} P_n(t, t + \Delta t) = o(\Delta t)$$

方能说明系统中顾客的到达形成了泊松流。

以上述前提条件为基础，为此可以建立多服务台模型(M/M/S/ ∞)：

(1) 系统服务台数量为 c ($c \geq 1$)，此处设定 $c=4$ 或 $c=5$ ；

(2) 设顾客单个到达，系统中顾客到达均符合泊松分布，到达率为 λ ；

(3) 系统中服务台的服务时间相互独立，服务时间服从参数为 μ 的负指数分布，并且平均服务率相同。当顾客到达时，若有空闲的服务台则该服务台马上为顾客提供服务，否则便排成一个队列进行等待，等待时间为无限；

(4) 系统中各随机变量相互独立且服务机制为先到先服务。

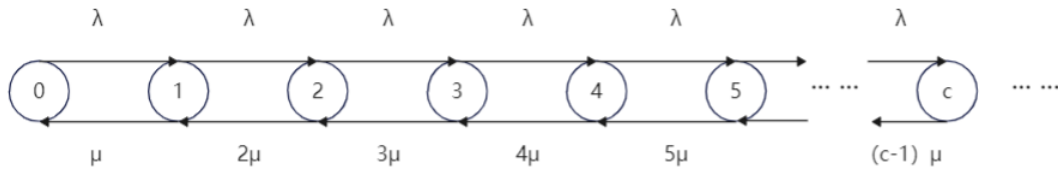


图2 模型的状态转移图

构建得系统状态概率的平衡方程如下：

$$\begin{cases} \mu \tilde{P}_1 = \lambda \tilde{P}_0 \\ (a+1)\mu \tilde{P}_{a+1} + \lambda \tilde{P}_{a-1} = (\lambda + a\mu) \tilde{P}_a, (a = 1, 2, 3, \dots, c-1) \\ c\mu \tilde{P}_{a+1} + \lambda \tilde{P}_{a-1} = (\lambda + c\mu) \tilde{P}_a, (a = c, c+1, \dots) \end{cases}$$

式中， ρ 是系统中至少有一个顾客的概率，也就是检测台处于忙的状态的概率，因而也称为服务强度，它反映了系统的繁忙程度。 $\sum_{a=1}^{\infty} \tilde{P}_a = 1$ 只有在 $\rho = \frac{\lambda}{c\mu} < 1$ 的条件下才能得到，即要求顾客的平均到达率小于系统的平均服务率，

才能使系统达到统计平衡。

稳态指标：

$$\begin{aligned}\tilde{P}_0 &= \left[\sum_{a=0}^{c-1} \frac{1}{a!} \left(\frac{\lambda}{\mu}\right)^a + \frac{1}{c!} \cdot \frac{1}{1-\rho} \cdot \left(\frac{\lambda}{\mu}\right)^c \right]^{-1} \\ \tilde{P}_a &= \begin{cases} \frac{1}{a!} \left(\frac{\lambda}{\mu}\right)^a \tilde{P}_0, & a \leq c \\ \frac{1}{c! c^{a-c}} \cdot \left(\frac{\lambda}{\mu}\right)^a \tilde{P}_0, & a > c \end{cases} \\ \tilde{L}_q &= \frac{(c\rho)^c \rho}{c! (1-\rho)^2} \tilde{P}_0; \quad \tilde{L}_s = \tilde{L}_q + \frac{\lambda}{\mu} \\ \tilde{W}_q &= \frac{\tilde{L}_q}{\lambda}\end{aligned}$$

式中， λ 表示顾客的到达率，即每分钟有 λ 人到达服务系统； c 表示服务台的个数； μ 表示平均服务率，即一个服务台可以服务 μ 人/分钟； \tilde{L}_q 表示系统平均等待队长； \tilde{L}_s 表示系统的平均队长； \tilde{W}_q 表示系统中顾客平均等待时间。

设 \tilde{L}_{qmin} 与 \tilde{L}_{qmax} 为系统平均等待队长的最小值与最大值； \tilde{W}_{qmin} 与 \tilde{W}_{qmax} 为系统平均等待时间的最小值与最大值。可设置服务强度 ρ 与检测点日容量的最优约束条件为：

$$\begin{cases} \tilde{L}_{qmin} \leq \frac{(c\rho)^c \rho}{c! (1-\rho)^2} \tilde{P}_0 \leq \tilde{L}_{qmax} \\ \tilde{W}_{qmin} \leq \frac{\tilde{L}_q}{\lambda} \leq \tilde{W}_{qmax} \end{cases}$$

并由此得到最优服务强度 ρ^* 与最优检测日容量。

5.1.3 核酸混测模型的建立

考虑实际情况，在设置核酸混检方案时应综合考虑核酸混合检测平均等待队长以及每组每人的平均检测次数。因此，结合上述离散型变量的数学期望模型与排队论模型，建立最优约束条件：

$$\begin{cases} \min f(x) = 1 + \frac{1}{x} - q^x \\ x \leq \frac{(c\rho)^c \rho}{c! (1-\rho)^2} \tilde{P}_0 \leq cx \end{cases}$$

使用 MATLAB 编程，寻找满足上述最优约束条件的正整数 x 。

5.2 问题二的模型建立与求解

5.2.1 数据与资料收集

在本轮疫情中，深圳和上海作为两个一线城市都承担着抗击疫情的重大责任。深圳作为防疫尖子生井然有序地应对着疫情难题，并对疫情进行了有效的控制，相反，上海的疫情却迟迟无法得到控制，因此本研究以深圳市和上海市的真实数据的对比为参考进行模型假设。

对 2022-01-17 至 2022-04-27 期间深圳、上海新冠疫情感染人数数据进行收集和分析，并参考国家卫健委发布的《关于印发区域新型冠状病毒核酸检测组织实施指南（第三版）的通知》中对核酸检测的总体要求进行数据假设，对问题一所建立的模型采取了 4 种不同的方案。

（一）政府对核酸检测的总体要求

在精准快速的流调排查、社区管控基础上，科学研判疫情传播风险，划定区域核酸检测范围，制定具体工作方案，加强组织管理，优化采、送、检、报流程，进一步提升核酸检测质量和效率。新冠肺炎疫情发生后，所在的设区市，包括城区常住人口 1000 万以上的超大城市，应当在 24 小时内完成划定范围的区域核酸检测任务。

核酸检测的质量控制要求

通过加强质量控制，最大限度减少假阴性或假阳性，杜绝出具虚假核酸检测报告的现象。国家药品监督管理局医疗器械技术评审中心发布的《2019 新型冠状病毒核酸检测试剂注册技术审评要点》对新型冠状病毒试剂盒的要求：在企业设定的最低检测限以上，应 100%能够检测。在最低检测限附近的也应保持 90-95%的检出率。

采样点设置

各设区市应当综合人口数量、地理交通、核酸检测机构分布等情况，科学规划采样点布局。按每日 6 小时的采样时间计算，可参考 2000-3000 人设置一个采样点，600-800 人设置一个采样台，每个采样点需 4-5 个采样台（各地可按照社区、街道、乡镇、农村、城郊实际人口数量统筹设定），原则上以小区为单位设置采样点。

采样组织

在网格化管理基础上，加强核酸采样的组织动员，明确负责组织动员工作的部门和人员，规范组织动员工作流程。通过提前分时段预约、通知及采样等，减少人员在短时间内聚集，力争群众排队时间不超过 20 分钟，避免交叉感染。

原则上，疫情发生后的前 3 日，每日开展 1 轮区域核酸检测，摸清社会面潜在风险；之后根据流调溯源情况、社区封控范围、核酸筛查结果等综合研判，确定后续的筛查范围和频次。若社会面风险未得到控制、风险点位不清，则继续每日开展 1 轮区域核酸检测。当社会面风险得到控制、风险点位基本摸清时，可根据病例分布情况，针对有关街道、社区、小区划分重点区域，对重点区域每日一检，非重点区域可视情况隔日一检或 3 日一检，并按要求在规定的时间内出具检测结果。逐步缩小核酸筛查范围，提高筛查精度，科学精准筛查。

（二）深圳市新冠疫情累计感染人数数据

1 月 7 日 12 时-1 月 9 日 24 时，在深圳全域组织开展了一轮全市域核酸检测筛查，累计完成核酸采样 2218.78 万人次，已 100%完成检测，除密接流调发现病例 3、在重点人群核酸检测筛查中发现病例 4 以外，其余结果均为阴性。

截止 4 月 27 日，近两个月来上海累计本土确诊病例 42942 例，累计本土无症状感染者 490886 例，其中，10528 例无症状感染者后来转归为确诊病例，无症状感染者占阳性人群的比例超 92%。

5.2.2 混检方案的解决

根据现有资料可知，核酸检测准确性极高，故本问仍不考虑检测结果中存在的“假阴性”情况。且在按每日工作时间为 6 小时计算时，检测点平均服务人数的参考范围为 2000-3000 人，每个检测点设立的服务台参考数量范围为 4-5 个。同时，在设置核酸混检方案时，应尽量控制排队时间不超过 20 分钟。通过计算收集到的数据，可以得出深圳市患病率 p 约为 0.001。

本问讲结合上述收集到的资料，通过分析深圳和上海的实际案例，设置不同情况下的合适的混检方案。

案例一：2022 年 1 月 7 日，深圳市主城区 2 天完成了 2218.78 万人次的全员核酸。

根据收集到的已有资料，现以深圳市主城区第一轮全员核酸检测为例，主城区约有 2218.78 万人进行核酸检测，设进行核酸检测的总人数为 N ，核酸检测共分三轮，从 1 月 7 日 12 时开始，至 1 月 9 日 24 时结束，历时两天。共设置了 s 个核酸检测点，则平均每个检测点服务人数为：

$$N_0 = \frac{N}{s}$$

顾客的到达率为：

$$\lambda = \frac{N_0}{T \cdot 60}$$

由问题一的分析可知，只有在 $\rho = \frac{\lambda}{c\mu} < 1$ 的情况下才能使系统达到统计平衡。

因为 $\rho = \frac{\lambda}{c\mu} < 1$ ，那么

$$\frac{N}{60 \cdot sc\mu T} < 1$$

可得：

$$\frac{N}{360 \cdot 4} < sc$$

代入已有数据可得平均每个检测点约 0.89 万人，系统平均服务率 $\mu=3$ 人/分钟。若要求一天内完成全员核酸检测任务，当日工作时间不超过 10 小时，则每个核酸检测点至少需要设置 4 核酸检测台，即 $c \geq 4$ 时，才能保证系统服务强度 $\rho < 1$ 。

在上述信息的基础上，控制日工作时间分别为 6 小时、5 小时和 4 小时，编程计算此时检测点预期人数与计划检测点数量，结果如下：

表 4 控制日工作时间 $T=6$ 系统指标分析

$T=6$					
c	2	3	4	5	6
检测点预期人数	2160	3240	4320	5400	6480

计划检测点数	3424.043	2282.695	1712.021	1369.617	1141.347
--------	----------	----------	----------	----------	----------

表 5 控制日工作时间 T=5 系统指标分析

T=5					
c	2	3	4	5	6
检测点预期人数	1800	2700	3600	4500	5400
计划检测点数	4108.851	2739.234	2054.425	1643.540	1369.617

表 6 控制日工作时间 T=4 系统指标分析

T=4					
c	2	3	4	5	6
检测点预期人数	1440	2160	2880	3600	4320
计划检测点数	5136.064	3424.043	2568.032	2054.425	1712.021

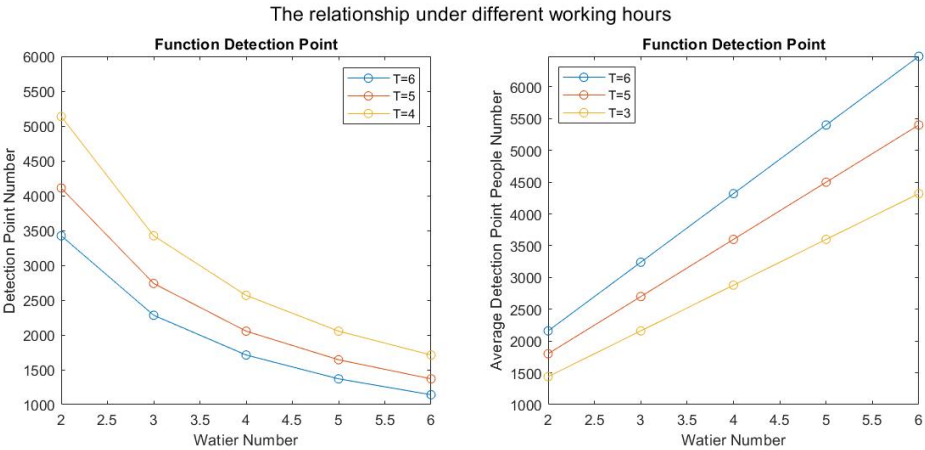


图 3 在不同的工作时间内服务台的个数与各变量间的关系

结合上述计算结果与收集到的数据进行考虑，当日工作时间为 6 小时，服务台数量为 4 或 5 个时，检测点预期人数符合预期，且计划检测点数量相对较少。故选择确定日工作时间为 6 小时，一个检测点设置 4~5 个检测台。

设进行核酸检测所需医护人员人数为 R ，控制日工作时间为 6 小时，分析 $C=4$ 和 $C=5$ 时系统的各项指标情况：

表 7 当日工作时间 T=6 时 c=4 与 c=5 的系统指标对比

c	T	λ	ρ	\tilde{L}_q	\tilde{L}_s	\tilde{W}_q
4	6	11.9305	0.9942	170.5786	174.5555	14.2977
5	6	14.8871	0.9925	130.9207	135.8831	8.7942

仍以深圳主城区全员核酸检测为例，要求两天内完成核酸检测任务，因受场地大小、病毒传播感染可能性、以及居民可接受的排队时间等因素的约束，

取 $\tilde{W}_{qmax}=2$ ，保持日工作时间为 6 小时，分析当服务台数量分别为 4 个或 5 个时系统各指标情况：

表 7 当顾客平均等待时间 $\tilde{W}_{qmax}=2$ 时 c=4 与 c=5 的系统指标对比

\tilde{W}_{qmax}	c	λ	ρ	\tilde{L}_q	\tilde{W}_q	s	q	R
2	4	11.5223	0.9602	23.1601	2.01	1783	4148	14264
2	5	14.5189	0.9679	29.2113	2.012	1415	5227	14150

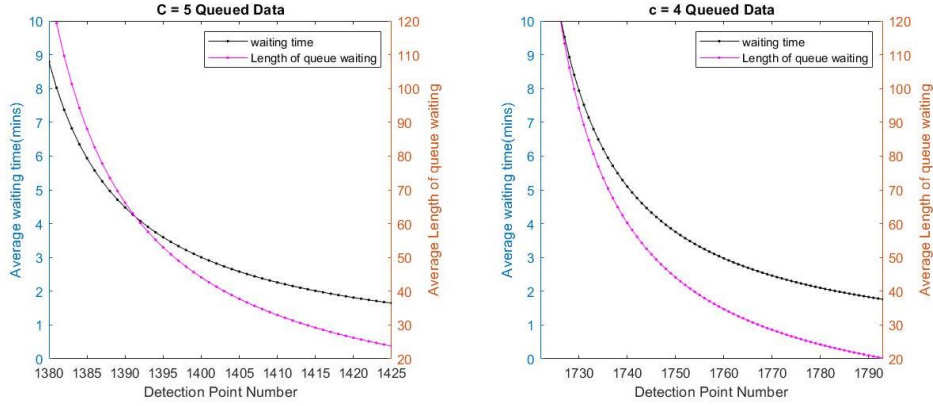


图4 在不同的服务台个数下检测点与平均等待时间的关系

结合上述两个表中的对比数据，以服务强度、系统平均等待队长、系统中顾客平均等待时间为评价指标。可以看出在工作时间为 6 小时的条件下，虽然 $c=4$ 时服务强度 $\rho=0.9942$ ，相较于 $c=5$ 时的 $\rho=0.9925$ 略大，但是 $c=4$ 时系统平均等待队长为 170.5786，系统中顾客平均等待时间为 14.2977，远大于 $c=5$ 时的系统平均等待队长 130.9207 和系统中顾客平均等待时间 8.7942。考虑到较长的排队时间可能会增大新冠病毒传播感染的风险，因而在控制日工作时间的情况下， $c=5$ 的情况会优于 $c=4$ 的情况。在取系统中顾客平均等待时间的最大值 $\tilde{W}_{qmax}=2$ 的条件下， $c=5$ 时的各项系统指标，包括服务强度 ρ 、检测点平均容量 q 和所需医护人员数量均优于 $c=4$ 的情况。

因此，综合评价下来，将选定日工作时间为 6 小时，服务台数量为 5 个。由此可以确定平均等待时间的最大值 $\tilde{W}_{qmax}=2$ ，进一步可以确定系统平均等待队长为 $\tilde{L}_q=29.2113$ 。

通过代入第一问设置的服务强度 ρ 与检测点日容量的最优约束条件：

$$\tilde{L}_{qmin} \leq \frac{(c\rho)^c \rho}{c!(1-\rho)^2} \tilde{P}_0 \leq \tilde{L}_{qmax}$$

设置 \tilde{L}_{qmin} 为 k ， \tilde{L}_{qmax} 为 ck ，即可得到：

$$k \leq \frac{(c\rho)^c \rho}{c!(1-\rho)^2} \tilde{P}_0 \leq ck$$

代入求解可得 $4 \leq k \leq 29$ 。

在患病率 $p=0.001$ 的情况下，根据表 2 中的不同 p 值对应的最优 k 值以及期望 $E(X)$ 可以得出最优 k 值为 32。又由于 $4 \leq k \leq 29$ ，故取 $k=29$ ，即每组混检人数最优为 29 人。

案例二：2022 年 1 月 13 日 19 时，上海奶茶店的 5 例病例中，已累计排查到上述 5 人在沪密切接触者 304 人，均已落实隔离管控，新冠病毒核酸检测结果均为阴性。累计排查到在沪密接的密接 2830 人，均已落实隔离管控，新冠病毒核酸检测结果均为阴性。累计排查筛查人员 30783 人，其中 30128 人新冠病毒核酸检测结果均为阴性，其余正在检测中。

针对案例二中累计排查筛查人员，即可理解为低风险人群，由于其数量众多，为了更快的得到核酸检测结果，防止新冠病毒扩散，故此次核酸检测将采用快速、高效的原则，根据案例一中 2000~3000 人数法则得出的结论，选取日工作时间为 4 小时的条件，比较分析服务台数量 $c=3$ 与 $c=4$ 时的不同结果。

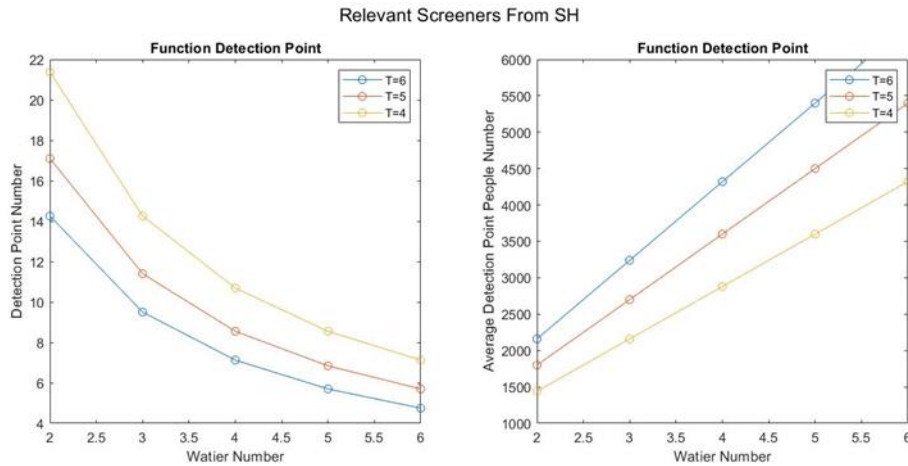


图 5 针对上海的案例，在不同日工作时间内各变量间的关系

表 8 当日工作时间 $T=4$ 时 $c=3$ 与 $c=4$ 的系统指标对比

T	c	λ	ρ	\tilde{L}_q	\tilde{W}_q	s	q	R
4	3	8.5508	0.9501	18.087	2.1152	16	2053	96
4	4	10.6885	0.8907	7.2594	0.6792	12	2566	96

因为在日工作时间 $T=4$ 的条件下，虽然 $c=3$ 时系统平均等待队长为 18.087，系统中顾客平均等待时间为 2.1152，大于 $c=4$ 时的系统平均等待队长 7.2594 和系统中顾客平均等待时间 0.6792，但是考虑到此时筛查人员被感染或携带新冠病毒的可能性较低，故这两项并不作为此时的主要参考指标。这时，注意到 $c=3$ 时服务强度为 0.9501，远大于 $c=4$ 时的 0.8907，由于针对这类人群的核酸检测主要注重高效、快速，因此以服务强度为主要参考指标，再考虑到 $c=3$ 或 $c=4$ 时所需的医护人员数量相同，故选定服务台数量 $c=3$ 。

由此确定平均等待时间的最大值 $\tilde{W}_{qmax}=2.1152$ ，系统平均等待队长为 $\tilde{L}_q=18.087$ 。同理，通过代入第一问设置的服务强度 ρ 与检测点日容量的最优约束条件：

$$\tilde{L}_{qmin} \leq \frac{(c\rho)^c \rho}{c!(1-\rho)^2} \tilde{P}_0 \leq \tilde{L}_{qmax}$$

设置 \tilde{L}_{qmin} 为 k , \tilde{L}_{qmax} 为 ck , 即可得到:

$$k \leq \frac{(c\rho)^c \rho}{c!(1-\rho)^2} \tilde{P}_0 \leq ck$$

代入得 $k \leq 18.087 \leq 3k$, 求解可得 $6 \leq k \leq 18$ 。

将患病率 $p=0.001$ 代入表 2 不同 p 值对应的最优 k 值以及期望 $E(X)$, 可以得出最优 k 值为 32。又因为 $6 \leq k \leq 18$, 故 $k=18$, 即每组混检人数最优为 18 人。

针对案例二中的次密接人群 2830 人, 由于次密接的情况较为严重, 为了防止疫情扩散, 这时 T 要尽可能的小, 根据下图数据假定 $T=2$, 比较分析服务台数量 $c=4$ 与 $c=5$ 时的不同结果:

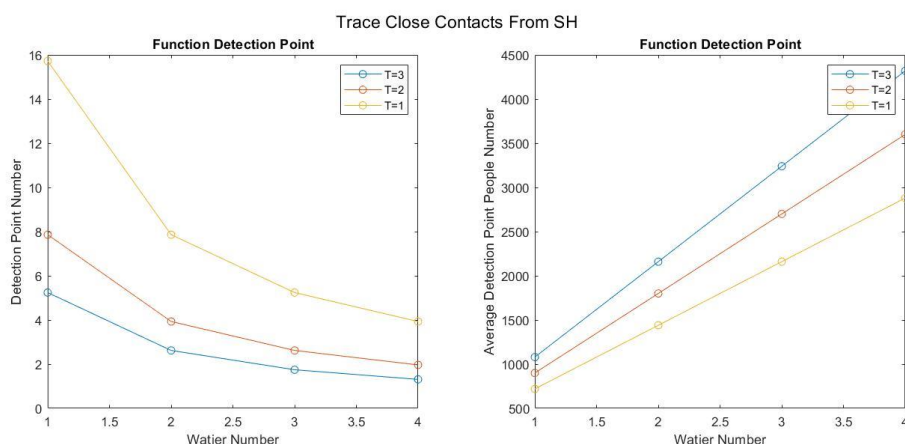


图 6 针对上海的案例, 在不同日工作时间内各变量间的关系

表 9 当日工作时间 $T=4$ 时 $c=3$ 与 $c=4$ 的系统指标对比

T	c	λ	ρ	\tilde{L}_q	\tilde{W}_q	s	q	R
2	4	7.8611	0.6551	1.2442	0.1583	16	2053	96
2	5	11.7917	0.7861	2.8892	0.6792	12	2566	96

在日工作时间 $T=2$ 的条件下, $c=4$ 时服务强度为 0.6551 不如 $c=5$ 时的 $\rho=0.7861$, 且二者系统平均等待队长和系统中顾客平均等待时间相差不大, 再考虑到服务台数量 $c=4$ 和 $c=5$ 时所需的医护人员数量相同, 故选定服务台数量 $c=5$ 。

此时系统平均等待队长 $\tilde{L}_q=18.087$ 。同理, 通过代入第一问设置的服务强度 ρ 与检测点日容量的最优约束条件, 并设置 $\tilde{L}_{qmax}=ck$, 这时由于 \tilde{L}_q 足够小, 故不考虑不等式左边的情况, 因此可得到:

$$\frac{(c\rho)^c \rho}{c!(1-\rho)^2} \tilde{P}_0 \leq ck$$

代入数据可得 $2.8892 \leq 2k$, 即 $k \geq 2$ 。此时取 $p=0.01$, 根据表 2 中的数据

可得 $k=11$ ，满足 $k \geq 2$ 的条件，故取 $k=11$ ，即每组混检人数最优为 11 人。

针对案例二中的 304 位密切接触者，假设密接都被隔离，因此不需要考虑排队问题。由于密切接触者被感染或携带病毒的概率极高，由问题一中设立的离散型随机变量的数学期望模型可知，当 $p \geq 0.1$ 时，应采用单人单管的核酸检测方法，即 $k=1$ 。故在此情况下，选取 $k=1$ 。

5.3 问题三的模型建立与求解

5.3.1 模型的假设

根据问题二中收集到的国家卫健委发布的文件，再借鉴深圳市高效的防疫政策，针对问题三的模型，提出以下假设：

- (1) 疫区所有人全部进行居家隔离，足不出户，每个人相对独立；
- (2) 每个小区互相独立，无接触；
- (3) 核酸检测时，被测者的到达呈均匀分布，有秩序地排队，不会出现排队过长。排队时戴好口罩，保持一米以上的间隔，保证做核酸时新冠病毒不会传播；
- (4) 一旦检测结果呈阳性，无论是无症状还是确诊患者，都立刻被隔离。

5.3.2 期望模型的建立

设置核酸检测错误率为 r ，包括由于采集病毒数量不足，操作不当等流程中存在的所有可能导致核酸检测有误的概率，所以通常我们取 $[0.1 \ 0.3]$ ，为了使数据更为立体，我们在计算过程中也加入了 0.01, 0.02, 0.05 等数据。

新冠病毒患病率 p_1 ，样本检测结果为阴性的概率 $q = 1 - p$ 。由此可知，在第二轮核酸检测中患病率为 $p_2 = p_1 r$ 。

设 X_i 分别表示在第 i 轮核酸检测中每组里每个人需要的检测次数， k_i 表示第 i 轮核酸检测中每组混检人数。

$$X_i = \begin{cases} \frac{1}{k_i}, & \text{该组检测结果为阴性} \\ \frac{1}{k_i} + 1, & \text{该组检测结果为阳性} \end{cases}$$

由此易得出 X_i 的分布律为：

$$P(X_i = \frac{1}{k_i}) = q_i^{k_i}; \quad P(X_i = \frac{1}{k_i} + 1) = 1 - q_i^{k_i}$$

根据第一问可知，随机变量 X 的数学期望 $E(X)$ 就能够表示每组中每个人的平均检测次数： $E(X) = \frac{1}{k} \cdot q^k + (1 + \frac{1}{k})(1 - q^k)$ ，化简可得 $E(X) = 1 + \frac{1}{k} - q^k$

因此，

$$E(X_i) = 1 + \frac{1}{k_i} - q^{k_i}$$

在多层核酸混合检测中，为了有效降低检测工作量，应取总体核酸检测次数的最小值，根据本文建立的模型，故使用多目标规划，设定如下目标函数：

$$\min\{E(X) = \sum_{i=1}^n E(X_i)\}$$

设置指标 α 来衡量核酸检测错误率 r , $1 \leq \alpha \leq n+1$ 。初始状态下 $\alpha = 1$, 当 $k_i > 29$ 时, 称此时过载超过系统极限, 则 $\alpha = \alpha + 1$ 。指标 α 越小, 证明系统过载的越少, 即系统的核酸混检方案越合理。

以轮数作为横坐标, 以患病率为 $[0 \ 0.3]$ 为基准计算 α 的平均值作为纵坐标, 当取不同的 r 值时, 如图 8。

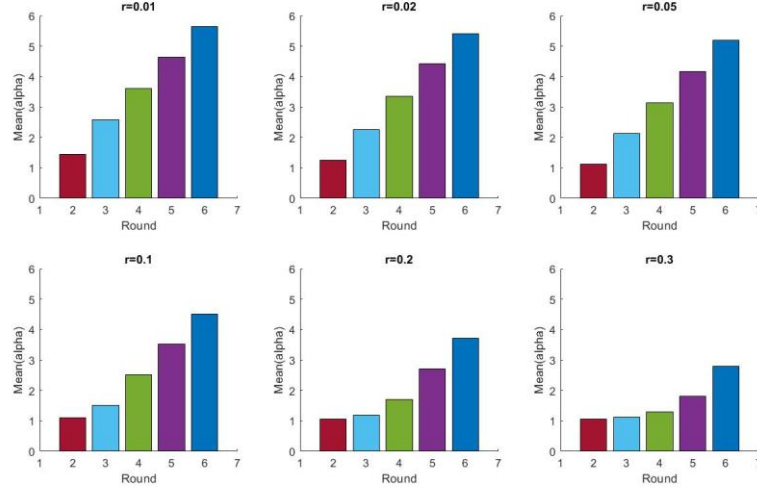


图 7 在不同核酸检测错误率 r 的情况下各轮检测中 α 平均值

从图 7 可以看出, 当 r 取 $[0.1 \ 0.3]$ 时, 核酸轮数 n 取 2 或 3 时, α 平均值明显比轮数 n 取 4, 5, 6 更为合理, 又因为通常多轮混检持续 7 天以上, 所以我们偏向于取 n 的最大值, 故决定取核酸轮数为 3, 即 $n=3$ 。

收集到相关资料显示, 2022 年 3 月 14 日至 20 日深圳封城一共 7 天, 按照“应检必检, 不落一人”的要求, 在封城 7 天内做了 3 轮核酸, 正好和文中取 $n=3$ 相照应。

已有文献资料发现, 带有新冠肺炎病毒的样本中, 每毫升通常含有百万以上个病毒。其他多份正常样本混合之后, 新冠肺炎病毒仅仅是浓度被稀释了, 病毒依然存在, 而且病毒均匀分布在样本保存液中。通常情况下, 在混合样本数小于 30 时, 混合样本后虽然病毒浓度被稀释了, 但对检测结果影响很小。故此处设立约束条件:

$$0 < k_1, k_2, k_3 < 30$$

因此, 建立模型

$$\begin{cases} \min E(x) = \sum_{i=1}^3 E(X_i) \\ 0 < k_1, k_2, k_3 < 30 \end{cases}$$

当 r 取不同的值时, 拟合出患病率 p 和 k_1, k_2, k_3 的以及期望最小值的函数关系, 如图 8. 图 9.

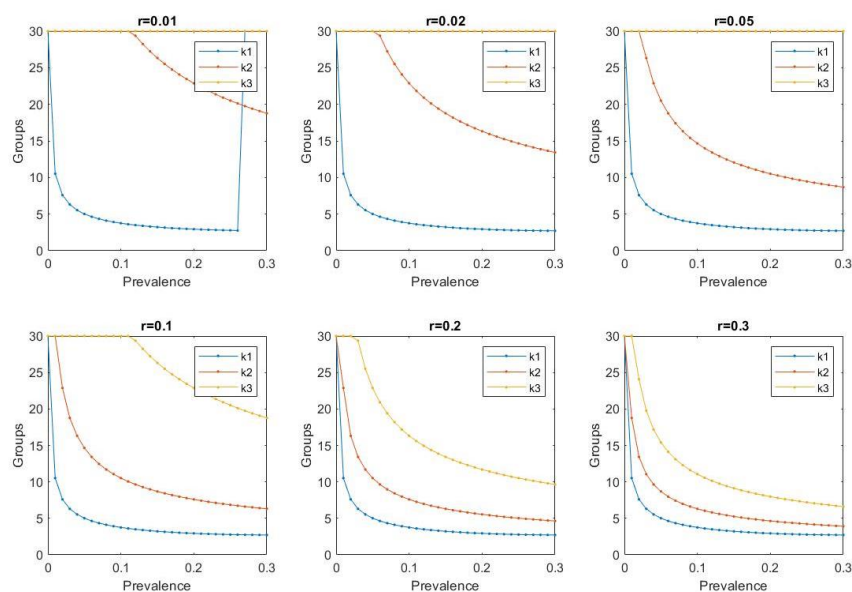


图 8 在不同核酸检测错误率 r 的情况下各轮检测中 k 值的最优取值

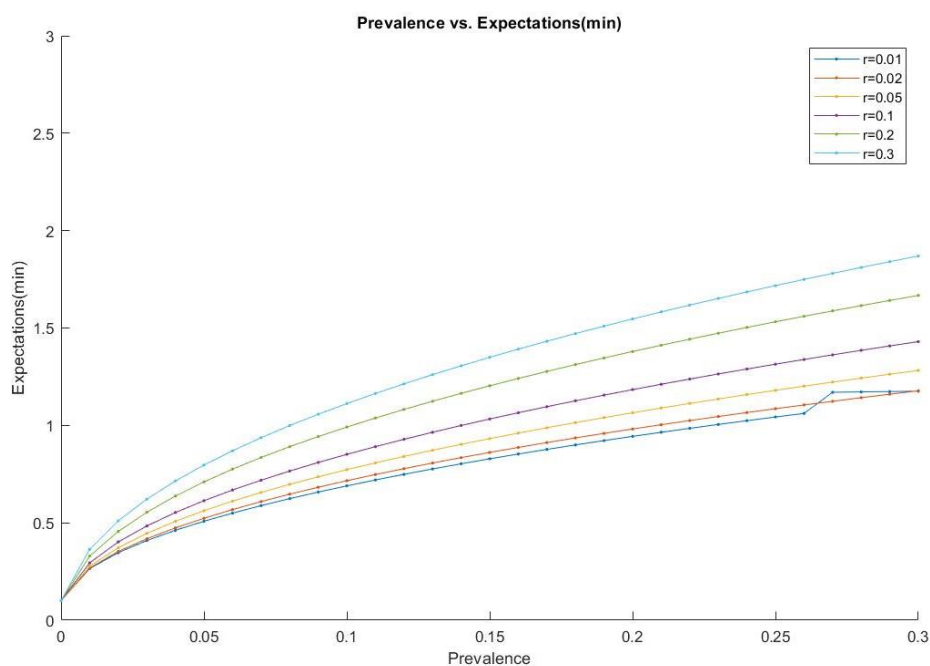


图 9 不同错误率 r 的情况下期望和患病率 p 的关系

由此, 若是得知需要多轮混检城市的错误率 r 和患病率 p , 就可以求得 $E(X)$ 的最小值和每轮 k_1 、 k_2 、 k_3 的值, 确定在多轮核酸检测的情况下最优的混检方

案。如表 11(从左到右依次为 k_3 、 k_2 、 k_1)给出了几个代表值。

表 11 给定条件下 k_1 、 k_2 、 k_3 的值最小值

$r \backslash p$	0.1	0.05	0.01
0.2	17, 8, 4	23, 11, 6	30, 23, 11
0.3	12, 7, 4	16, 9, 6	30, 19, 11

6、模型的分析与评价

6.1 模型的评价

6.1.1 模型的优点

(1) 本文运用 MATLAB 对建立新冠病毒核酸混合采样检测模型进行运算，具有较强的说服力，并且较好的解决了数据体量巨大的问题。

(2) 本文借助排队论模型，考虑了核酸检测过程中的排队现象，使得设置的核酸混合采样检测方案更贴近现实生活，且更容易被大众所接受。

(3) 本文考虑了现实生活中核酸检测过程中存在的新冠病毒传播感染问题，模型更合理。

6.1.2 模型的缺点

(1) 在一般情况下，某个确定的疾病筛查区域，由于防疫政策等各种因素，患病率 p 是会随着时间不断变化的，在进行全体的核酸检测筛查之前 p 不是常数，利用数据统计的方法，通过历史的累计确诊人数以及总人口数确定出的患病率仅为当时该地区患病率 p 的近似值。

(2) 实际情况中，每组混合采样检测人数 k 还应与核酸检测的准确性有关，问题一中的离散型随机变量的数学期望模型未考虑“假阳性”、“假阴性”的情况。

(3) 在问题三的模型中，患病率 p 与错误率 r 应该存在某种关系，但在此问的模型建立过程中未考虑这一关系。

6.2 模型的推广

本文所建立的离散型变量的数学期望模型、排队论模型不仅考虑到了现实生活中进行核酸检测时的排队过程和等候过程中病毒的传染过程，使得模型更加贴合现实，并且模型还能推广至其他传染病的检测问题研究中去。运用历史数据对模型参数进行估计，这对于核酸混合采样检测方案的设立具有参考意义。

7、小结

在本轮疫情中，上海防疫太依赖流调，但是流调失利之后，又缺乏应急预案，这就导致防疫兵败如山倒。上海防疫失控的主要问题包括对城市防疫管理系统的不够重视；基层动员能力不足，没有做好应急准备和供应链组织存在严重问题。

但是，上海作为一线大城市，在此轮防疫中也有值得肯定的地方：

- (1) 上海为全国承担了 35%以上国际航班外来输入型压力；
- (2) 上海精准防疫“以最小的代价实现最大的防疫成果”。

本文通过分析深圳与上海的真实案例，希望能给上海核酸混检方案提供一定参考意见，祝上海早日战胜疫情，恢复这座城市昔日的活力。

8、参考文献

- [1] 杨涵莽, 王少峡. 新型冠状病毒核酸检测技术研究进展[J]. 包头医学院学报, 2022, 38(02): 77-82. DOI:10.16833/j.cnki.jbmc.2022.02.017.
- [2] 嵇冉, 王健. 巨大量核酸检测是如何完成的[J]. 中国统计, 2020(09): 19-21.
- [3] 赵小艳. 基于数学期望的新冠肺炎核酸检测方法[J]. 大学数学, 2020, 36(06): 19-22.
- [4] 凌婷婷, 丁伯伦, 段双双. 基于排队论的全员核酸检测点优化设置[J]. 太原学院学报(自然科学版), 2022, 40(01): 65-68. DOI:10.14152/j.cnki.2096-191X.2022.01.011.
- [5] 中华人民共和国国家卫生健康委员会. 《关于印发区域新型冠状病毒核酸检测组织实施指南(第三版)的通知》[Z]. 1991—06—05