首战数学建模，三天三夜也没有那么难

**背景回顾**

数学建模比赛多种多样，五一数学建模虽说不是含金量最高的比赛，但总的来说它还是有一定知名度和影响力的，而且它的题目难度没有高教社杯那么难，尤其对于我们这种入门不久的选手来说（学了5周算是入门了吧），五一数学建模刚好可以作为我们的磨刀石，而且学院选拔种子队的首要依据就是拿到五一数学建模的成功参赛证，综合这些因素，五一数学建模肯定是我们的必选项了。

**心路历程**

**跃跃欲试的开始**

5月1号早上九点开题，头一天晚上我就开始紧张起来了，和潘总商量着明天早上几点起床、吃完饭几点能到教室、哪个教室环境好、哪个位置空间大还不热、要不要买一些吃的食物面包什么的，乱七八糟的想了一大堆，其实大部分考虑都没什么用，第二天早上我六点二十就醒了，怎么也睡不着了，🍞什么的也没有买，想着正餐还是必须要吃，零食面包什么的还是等等再买吧，吃完饭到西楼还不到八点，和潘总综合考虑打量以后还是选择了517教室，主要是519教室看着有点暗，桌子太大，也不是主教室，里面还有一些考研同学的📕什么的，我还担心后面会不会又有人来，最终还是选517了，后排靠窗位置不错，空间大，视野好，还有一个窗台可以放书，这就很nice。后来也证明我们的确来早了，快九点了其它队伍才陆陆续续的过来。九点开题后大致浏览了一下三道题，A题是疫苗生产问题，B题是消防救援问题，C题是数据驱动的异常检测与预警问题，初步感觉A题像是排队论问题，B题是图论的路径类问题，C题应该是大数据分析问题，我们又仔细研究了一下题目，A题第一问很简单，计算数据属性的，平均值、方差什么的，但后面的题目就看不大懂了，题目的含义都没搞明白，再看B题，前三题都是预测拟合类问题，处理数据什么的，感觉灰色预测、曲线拟合、数据插值什么的应该都能用，后面选址问题求一下最短路径估计就行了，C题应该是要分析出一个指标来判断生产线是否出现了异常，联想起之前老师说过C题看起来简单，但统计指标的选取非常重要，一旦统计指标选错了后面就会全错。最终，我们还是还是选择了看起来我们知道模型最多的B题（其实这时想到的模型基本都用不了，主要还是经验少），选好题大概用了一个小时，联想国赛有人大半天都选不好题，我们就感觉自己的速度很快了，接着就正式开始写B题了。

**心力憔悴的过程**

首先我们查了大半个小时的论文，其实我们也没查到什么，总之乱七八糟的我们就开始写第一题了，第一题是分析春夏秋冬第一天怎么安排值班人数，阳姐说应该是规划类问题（我也觉得有道理），一天不超过30人，三个时间段最少不少于5人，但我又觉得这个问题我们不知道求值的那个方程，应该用预测方法来做，由于对灰色预测理解得不够透彻，这个题一开始我们竟然想用灰色预测来做，但由于数据实在是太少了，最后只能用权值占比来做了，但我老感觉明年这四天各时间段发生事件的概率和前四年的关系不是很大，但也没有其它好的方法了，这个题拖到最后一天晚上还是用权值占比的方法写了，第二题是要求用前四年的数据建立数学模型来预测第六年各月事件的次数，用第五年的已知数据进行检验，这个题数据很多，一开始我觉得这个肯定是用灰色预测了，没想到千辛万苦预测出来竟然是一条直线？？？这时我就有点崩溃了，连续两题都没写出来，还是前两题，我还是不死心的用曲线拟合来拟合48个散点，由于离散程度太高，拟合效果非常差，挣扎一段时间后，我们终于观察到48个散点呈周期性变化且以12为周期，这是一个非常重要的信息，于是我们开始找周期性预测的模型，但我们又错误的选择了蒙特卡洛，这是一个随机预测模型，即对下一次事件的预测只与当前状况有关（这明显不符合），当时决定用蒙特卡洛模型后我开始在CSDN上找源码，但有些源码文件只有会员才能下载，会员费用要50，而且也不能确定下的代码是否正确，下午的时间就在找源码的过程中结束了，最终在吃晚饭的时候我看到MATLAB交流群里的信息，淘宝5毛钱可以下载一次CSDN代码，于是我回去后又来精神了，开始在淘宝、CSDN上来回折腾，结果找到的代码要么是不对、要么是不能运行报错、要么是版本过老2020a不支持，而且在搜索的过程中我们终于也发现了蒙特卡洛模型的弊端，第2题根本不能用蒙特卡洛来做。这个时候我已经有点想换题了，主要下午听到隔壁已经在写B题第四问了，其它3支队伍进度好像都比我们快很多，而我们前两问一个都没做出了，这确实对我是一个不小的打击，但总算还是忍住了，如果这个时候换题，估计我们最后论文都要交不上了（事实证明坚持下去是对的），然后在不断的挣扎中终于找到正主，时间序列模型闪亮登场，但遗憾的是我们依然没有这个模型的源码，《数学建模算法与应用》这本书上其实是有时间序列模型的源码的，只可惜的是我买的是第3版的，4月份刚好买到了发行的新版，于是我们又陷入了困境，然而功夫不负有心人，我又在之前加的另一个群里找到了第2版的源码文件，本以为终于要解决这道问题时发现用时间序列模型直接预测结果明显偏大，这时已经11点，为了第二天有充足的精力与时间完成剩下的题，我们决定先回去睡觉💤，第二天直接开始第三题。第三题是建立各类事件发生次数与月份关系的多种数学模型，以拟合度最优为评价标准，确定每类事件发生次数的最优模型。这道题还是一头雾水，查了一上午资料，最后还是在午饭之前用曲线拟合把这道题过了，下午就是第四、第五题了，这两道题思路还是很清晰的，没有走多少弯路，主要就是在理解题意上花了不少时间，三维图画出来后在图上就可以看出很多信息，这些信息整理一下就可以直接拿来当结论了。综合这几道题的解题过程，我觉得拿到题后一定不能先急着确定模型，一定要先对数据有一个总体上的直观把握，最好的方式就是先画图，观察数据的规律趋势后再选定模型。大概到了晚上我们就把四五题给写完了，第六题思路也很流畅，这道题是选址问题，首先肯定要得到各区域之间的最短路径，因为这关系到救援到达的时间，一开始打算用dfs搜一下，后来发现图论的Floyd 算法时间复杂度更低、名字听起来也高大上一些，于是决定用Floyd算法，我们还确定了已有消防站对各区域的救援程度、人口密度、事件发生次数等主要影响因子，至此今天的工作就全部完成了，感觉有点飘，主要题目差不多都写完了，第六题也有了明确的思路，回到寝室洗了个澡，用Java把dfs简单实现一下大概就12点半了，睡觉💤。第三天上午就是第六题的建模过程了，基本数据都准备好了，我们计划用贡献值代表在某个区域建消防站的贡献程度，最后贡献值最大那个地域就是建消防站的位置。首先要算出每个消防站点贡献值，本以为思路都明确了应该很快就能做出来的，结果在求贡献值的过程中又发现了一些错误，首先是路径影响因子的计算，一开始我们是用每个区域到其它所有区域的和来作为路径影响因子，再用路径影响因子乘其它所有因子后求和作为贡献值，但其实这样是错的，因为其余每个区域的事件密度、人口密度不同，到达每个区域的最短路径也不同，所以应该分别计算到各个区域的最短路径再点乘其它影响因子后再求和，这个结果才是一个区域真正的贡献值。但这个贡献值计算有一个难点，贡献值的含义是该区域对其它所有区域的救援帮助能力之和，很明显，对某个区域的救援帮助能力应该和到这个区域的最短距离呈负相关，即距离越远帮助越小，所以我们计划用最短路径的反比来代表距离影响因子，这是一个反比例函数，问题就出在我们无法计算出它对自己所在区域的贡献值，因为我们假设所有区域都是一个物理上的点，这样到自己区域的距离就是零了，反比例函数在x=0处的值是无限大，肯定不能用无限大代表对自己所在区域的贡献值，也不能随便找一个常数来代替，我们应该用一个函数来确定这个距离影响因子，用常数来代替就相当于我们在一个因子的计算上用了两个标准，这肯定是不行的，当时我一直没想到怎么量化对自身贡献的大小，队友也坚定说不用考虑对自身的贡献，我们在这个问题上激烈讨论了很久也没有得到一个统一的结论，最终我们还是选择不考虑对自身贡献值的影响，最后得到的答案是D点，我也觉得这个答案是对的，但我还是感觉我们得到的方式不对，如果考虑对自身贡献的影响，结果肯定是P点，但考虑到P点人口密度过高、占地面积过小且还是交通枢纽，我们应该舍弃P点转而将消防站建立在贡献值第二高的D点，其实这也不难实现，我们原来是用反比例函数来计算距离影响因子的，但在x轴上单调递减且在x=0处不为无穷大的函数还有很多，比如对数函数，如果当时能想到用对数函数来计算距离影响因子的话结果肯定会好很多。第六题搞完已经到下午了，但这时离结束还早，由于在解题过程中我一直忽视了向写论文的队友✍传达我们的建模思路，我原本以为稍微修改一下语句理一下逻辑就好了，但结果发现论文基本上没有逻辑，而且论文里有很多专业性内容，有专业性内容当然很好，但主要是这些内容没有串联到我们的建模思路中，简单的来说就是我们根本不会用这些知识，这样就导致整篇论文看起来华而不实、完全没有逻辑。我觉得，评委们虽说每篇文章看得时间短，但他们肯定是能看懂推导过程的，如果说我们根本没有推导过程，只是用看起来高大上的公式内容堆砌起来，我觉得我们会死得明明白白的……，不说了，说多了都是泪，主要还是我们的经验不足。然后就是改论文了，先把逻辑不通顺、用不到的知识删掉，再把自己的思路加进去，摘要、优缺点、推广、假设啪啪啪一敲，史上最强论文就这样诞生，没有之一，上传提交，直接起飞✈。  
（PS：其实图号和表号还没调，不过这些都是我第二天发现的了）

**凌晨两点的收尾**

五月四号凌晨两点半🕖，门卫大爷已经在大厅椅子上睡着了，走出和平大楼🌃，外面已经下起了小雨🌧，刚才改论文的时候就听见窗外一直在刮风，伸个懒腰，现在我们只想回去睡个通天亮💤💤。

**写在最后**

**结束，也是新的开始**

五一建模只是我们评选种子队的一场测试，国赛才是我们的目标，距离国赛也仅有4个多月的时间了，4个月后的结果又会是怎么样的呢？每个故事都会结束，只是有人欢喜有人哭。  
？？？  
打住，我又开始煽情了，先给国赛留个🕳吧，等国赛结束了再写篇国赛回顾，就这样，over。

2021年5月4日晚于锦城白宫