登峰杯

全国中学生学术科技创新大赛

浅谈投标游戏中制胜策略

北京市朝阳外国语学校
高一（1）班
林左 权衡 谭世茵 王悦晗

目录

1. 前言：问题的重述与构想----------3

2. 问题的解答----------3

2．1．问题的分析----------3

2．2．定义与符号的说明----------4

2．3．问题的模型假设----------5

2．4．问题模型的建立与求解----------5

2．4．1．假设分析与他组准则----------5

2．4．2．最终决策总流程----------6

2．4．3．流程制定的方法----------7

2．5．模型的评价----------10

2．5．1 模型的检验----------10

2．5．2 模型的优缺点----------11

3. 参考文献----------11

4. 附录----------11

登峰杯
全国中学生学术科技创新大赛

摘要

本题目源于 2016 年第一届“登峰杯”全国中学生数学建模竞赛，参赛期间我们针对这一投标游戏建立了数学模型，比赛后我们对这一问题进行了再分析，写出了这篇论文．这一数学问题可以看作仅清楚整体信息、决策目标，而不清楚每一个个体信息和决策集的多人博弈问题．本文从信息和物理中熵的角度出发，得出他组最终决策结果为正态分布，并通过分析求出这一分布中的多个参数，运用计算机模拟排除了一些非理性决策者和刻意干扰者的影响，得出我们的决策方法

关键词

博弈 不完整信息 非理性决策 计算机模拟 正态分布 最大熵原理

1. 前言：问题的重述与构想

问题请参考文末的附录．本文针对此问题的核心思想是从信息和物理中熵的角度出发，得出他组最终决策结果为正态分布，从而计算出最有利于自己的投标价．

2. 问题的解答

2．1．问题的分析

第一阶段投标之前，我们对一切实际投标游戏的情况没有任何了解与参照，于是我们从本组成员的心态出发，推己及人，推测其余各组的投标策略．

我们第一阶段采用如下的投标策略，并假设其余各组跟我们的策略相同：

1. 首轮投标的投标价 x_1尽量大，使自己第二轮投标时选择余地尽量大．但如果所得初始资金过少，则选择首轮投标做捣乱者、尽量干扰他组的判断，用余额购买信息．
2. 根据所得的首轮投标平均值$\overline{x_1}$确定有利于得分区间 A 的中轴 μ'和区间宽度 d'：其中，$\mu' \in [90,100]$，表达式见下文；$d' \in [2．5,5]$，随$\overline{x_1}$偏离 100 的程度增大而减小，其表达式略．
3. 若有余额，则视实际情况优先购买中位数、方差这两个信息．
4. 最终结合所购买的信息，根据实际情况，在所求得的区间内确定最终的投标值 x．

第一阶段投标结束后，我们统计了其结果和我们在几局中购买的信息，列出了如下的表格，通过分析它总结了各组投标的规律和我们组第一阶段投标策略的优缺点．

登峰杯
全国中学生学术科技创新大赛

第一阶段游戏情况										
局数	财富值 M	第一轮x_1	第二轮x_2	本组两轮平均值 x	总平均值. $\overline{X_1}$	区间中轴 μ’	区间范围 d’	得分	是否有 U 盘	可能的名次
1	97	97	97	97	97．091	97．303	4．94711	2	0	9，10
2	113	113	96	104．5	102．826	102．63	4．95009	0	1	1，2，3
3	93	93	92	92．5	92．261	93．758	4．62567	0	0	未知
4	100	100	91	95．5	95．609	96．091	4．87949	2	0	9，10
5	85	15	65	40	86．783	91．15	3．90819	0	0	未知

第二局的中位数和标准差分别是 104 和 9．637．第五局的中位数、标准差、最大值、最小值分别是 99,33．266,1106．63,120,0．

观察表格，我们发现：

1. 前 4 局的投标中，$\overline{x_1}$维持在较高水平，这可以说明大多数组的首轮投标价 $x_1 \approx$ 初始财富值 M，也就是我们在第一阶段中对各组 x_1尽量高的假设基本合理．我们自己也是如此投标，可以看出来这种策略在第一局和第三局中一定程度上“挽救”了自己．
2. 第 5 局中$\overline{x_1}$较之前偏低，我们分析一下原因：首先，我们自己抽到的初始财富值$M = 85$，按照以上假设就处在不利的位置．我们预计自己无望排到 4 至 10 名次，于是想办法阻碍别人得分．我们充当了“捣乱者”，第一轮仅投标 15，这就能将$\overline{x_1}$拉低，从而给可能存在的拥有极佳决策方法的组带来混乱；第二轮购买所有信息便于以后的分析．实际上，从方差和最小值数据，我们估计有 3 到 4 个组和我们的想法相同．这就说明会存在初始财富值较低的组选择“捣乱”的策略，并且有些组会意识到这一点．
3. 我们的结果并不太理想．分析我们已知的可能的名次，发现我们得过前 3 名或 9,10 名．这可以说明我们的利于得分的投标价格区间设置得太宽了，因此我们要让它变窄．我们对前 4 局的$\overline{x_1}$、x和最终结果比较，发现中标的x集中在$\overline{x_1}$附近．

2．2．定义与符号的说明

（1）初始资金为 M ($M \in N^*; M \in [80,120]$)；

（2）两轮投标价为 x_1, x_2，首轮投标平均价为$\overline{x_1}$，本局平均投标价为 $x = \frac{x_1+x_2}{2}$；

（3）最终投标金额正态分布区间的数学期望为 μ；

（4） 预期的“捣乱者”数量为 n，本文“捣乱者”指因初始资金过低而无望得分，并且在首轮投标中投 0 的选手数量；

（5）进入有利于得分的正态分布区间的概率为 P；

（6）正态分布的标准差为σ；

登峰杯 全国中学生学术科技创新大赛

（7）可花钱购买的首轮投标标准差为σ_1

2．3．问题的模型假设

1. 假设抽取财富量的过程绝对公平；
2. 假设各组均尽力想要获得好成绩；
3. 假设其他组第一轮投标价格遵照下文所述准则；
4. 假设其他组第二轮投标价格遵照下文所述准则；
5. 假设其他组决策准确度和焦虑程度与时间变化无关.

2．4．问题模型的建立与求解

2．4．1．假设分析与他组准则

由于每个人在考虑问题时考虑的是参与投标的整体，而一个个体的缺失对其决策的影响是微乎其微的，所以我们将自己“置身事外”，对他们进行分析.

1. 首轮投标

我们认为，财富值在[80，90]的人有在首轮投标中拉低平均值的能力，我们称他们为有捣乱条件的人．他们进行捣乱的概率随着他们手中财富升高而降低．同时，在其他区间，仍存在由于非理性因素而导致的首轮投标价格低于自己的财富值，为了便于分析，我们将他们等效替代为部分捣乱者和部分投出全部财富的人.

我们认为，其他组在本局中的首轮投标准则与我们在技术报告一中所提相似，但进一步的思考，我们发现，事实上，由假设 1，在每局中存在有捣乱条件，即扰乱首轮投标信息的人，但有捣乱条件的人不一定选择进行捣乱．假设有捣乱条件的人理性，那么则通过下表（决策结果表示该指定一名有捣乱条件的人的结果），可以分析出选择不进行捣乱是更优的决策，但由于存在非理性思考，对其他队的干扰再结合随机模拟的结果，我们认为每局存在 0-3 名捣乱者.

	其他有捣乱条件的人		
指定一名有捣乱条件的人	决策选择	全部选择捣乱	存在不捣乱者
	捣乱	有不大的概率获得成功	较大概率失败
	不捣乱	有较大的概率获得成功	较大概率失败

2. 二轮投标

登峰杯
全国中学生学术科技创新大赛

我们设他人由各自算法计算所得的结果为x_i,出现概率为$f(x_i)$，由于我们的估计具有一定的不确定性，同时为了便于分析，我们将其连续化．设其概率密度函数为$p(x)$，则其熵为$H(X)=-\int f(x_i)lnf(x_i)dx$．由于每人能否取得得分较高的名次不但取决于他们算法的准确度以及他们的理性程度，同时也与具有一定的随机性，正如物理学中一个系统中的粒子，能否处于最合适的状态不仅仅取决于它的势能和动能，也有随机因素在其中的，而对于一个仅清楚一部分约束条件的系统来说，熵最大的系统是最“真实”、最符合实际的．所以对于我们这个问题来说，我们要使$H(X)$最大，同时满足以下几个约束条件：

$$\int_{-\infty}^{+\infty} p(x)dx = 1$$

$$\int_{-\infty}^{+\infty} p(x)xdx = g(\bar{x}) = \mu$$

$$\int_{-\infty}^{+\infty} (x-\mu)^2 p(x)dx = \delta^2$$

即其满足一定的数学期望和方差的分布．这里的数学期望是一轮平均值的函数，因为可以认为他人所使用的算法相似，且主要使用数据为平均数；方差只与一轮平均值有关，因为一轮平均值在一定程度上反映了财富的分布，反映了二轮投标中投标价的分布．
由变分法，我们得出其分布满足正态分布．

2．4．2．最终决策总流程

1. 抽取初始资金 M，$(M \in N^*; M \in [80,120])$;
2. 根据 M 的大小，决定投标策略：
   2.1. 若$80 \le M \le 85$，我们扮演“捣乱者”，首轮投 0，第二轮购买全部信息；
   2.2. 若$85 < M \le 120$，我们扮演“全投者”，首轮投出全部，随后如下操作;
3. 获得首轮投标平均值数据$\overline{x_1}$
4. 计算出最终投标金额正态分布的数学期望 h：

$$h = \begin{cases} \frac{\overline{x_1}^2}{40} - 4\overline{x_1} + 250 & (\overline{x_1} \in [80,100])\ (h \in [90,100]) \\ -\frac{\overline{x_1}^2}{40} + 6\overline{x_1} - 250 & (\overline{x_1} \in (100,120])\ (h \in (100,110]) \end{cases}$$

5. 比较数值$\overline{x_1}$与 h
   5．1．若$\overline{x_1} \le h-5$，认为本组无望得分，第二轮购买所有数据便于以后分析
   5．2．否则继续保持不放弃，如下计算
6. 计算预期“捣乱者”数量 n，$n = -0.23\overline{x_1} + 23$
7. 计算出进入有利于得分的区间的概率 P，$P = \frac{8}{23-n}$，

登峰杯
全国中学生学术科技创新大赛

8. 计算正态分布的标准差$\sigma = \frac{\overline{x_1}^2}{160} - 1.25\overline{x_1} + 63$

9. 计算$1 - P$，查表得到它对应的标准正态分布的随机变量 t

10. 根据$x_2 = 2\sigma t + 2h - x_1$求出$x_2$

11. 根据余额决定是否购买其他信息

2．4．3．流程制定的方法

1．抽取初始资金 M

2．根据 M 的大小，决定投标策略：

2．1．针对抽到$80 \leq M \leq 85$的首轮投标，结合第一阶段实际游戏的数据可知自己得分的几　　率不大，所以我们的目的在于尽量干扰他组判断，故首轮投 0，在第二轮时购买全部信息，便于模型检验与分析（后续步骤不适用于此情况）；

2．2．针对抽到$85 < M \leq 120$的首轮投标，自己得分的几率很大，为使在第二轮投标时选择余地最大，故首轮全部投出，随后如下操作;

3．获取首轮投标平均值数据$\overline{x_1}$：

结合第一阶段实际游戏经验，我们更改了模型，使$\overline{x_1}$在其中发挥的作用更显著．本文以$\overline{x_1}$为自变量，设定一个最终投标金额正态分布的区间，根据$\overline{x_1}$调节其数学期望 h、预期“捣乱者”数量n、标准差σ等值，最终确定最佳的投标金额x_2．上述各因变量与自变量$\overline{x_1}$的关系的表达式的推导详见下述各步骤．

4．计算出最终投标金额正态分布的数学期望h：

在技术报告 1 中，我们用二次函数建立了一个利于得分区间中轴 h 关于的函数．其中，$\overline{x_1}$越大，h 越大；$\overline{x_1}$越偏离基准平均值 100，h 的变化量越大但变化率越小．结合上述函数关系以及该函数的极值，我们计算出了其表达式：

$$h = \begin{cases} \frac{\overline{x_1}^2}{40} - 4\overline{x_1} + 250 & (\overline{x_1} \in [80,100])\ (h \in [90,100]) \\ -\frac{\overline{x_1}^2}{40} + 6\overline{x_1} - 250 & (\overline{x_1} \in (100,120])\ (h \in (100,110]) \end{cases}$$

其图像如下:

登峰杯
全国中学生学术科技创新大赛

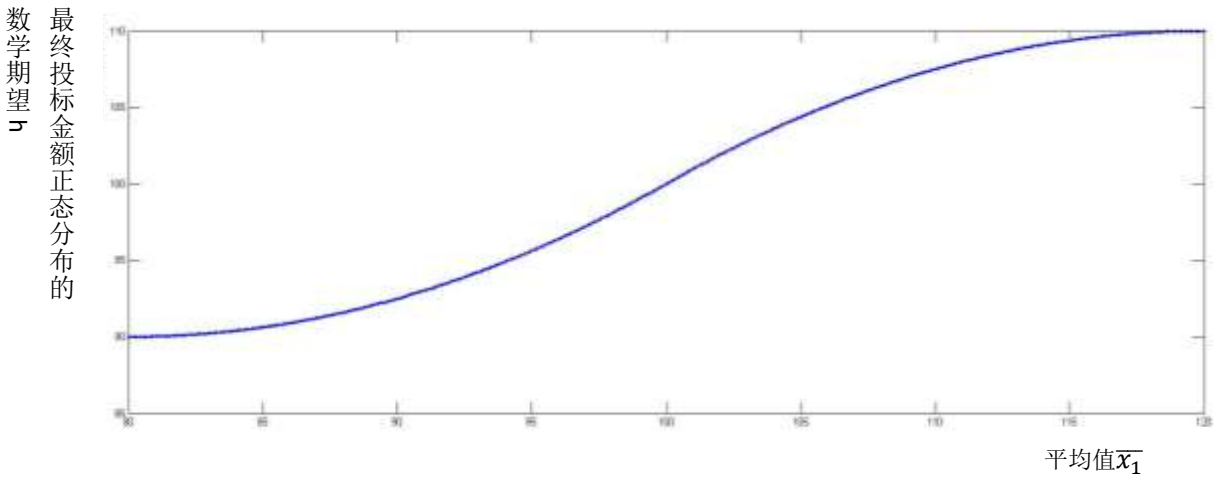


图 1——$h-\overline{x_1}$图像

经阶段一游戏实践检验，该模型能够较精确地预测出最终中标金额的分布情况: 在 1-4 局中我们能够根据现有数据推断出中标金额，经比较，中标金额与 h 的差距在±1 之内，故在此正态分布模型中采用 $\boldsymbol{h}$ 作为其数学期望.

5. 比较数值$\overline{x_1}$与 $\boldsymbol{h}$

5. 1. $\overline{x_1} \le h-5$时，即使第二轮全部投出尝试得分，但实际得分的几率已经很小，相比初始资金在 h 之上的选手我们的竞争力很弱，故这样决策很可能既无法获得分数，也无法获得有价值的信息. 反之，此情况下舍弃力争得分的希望，选择风险较小的购买信息方法就显得更为理智.（后续步骤不适用于此情况）

5. 2. $\overline{x_1} > h-5$时，继续保持不放弃，按照下述步骤继续进行.

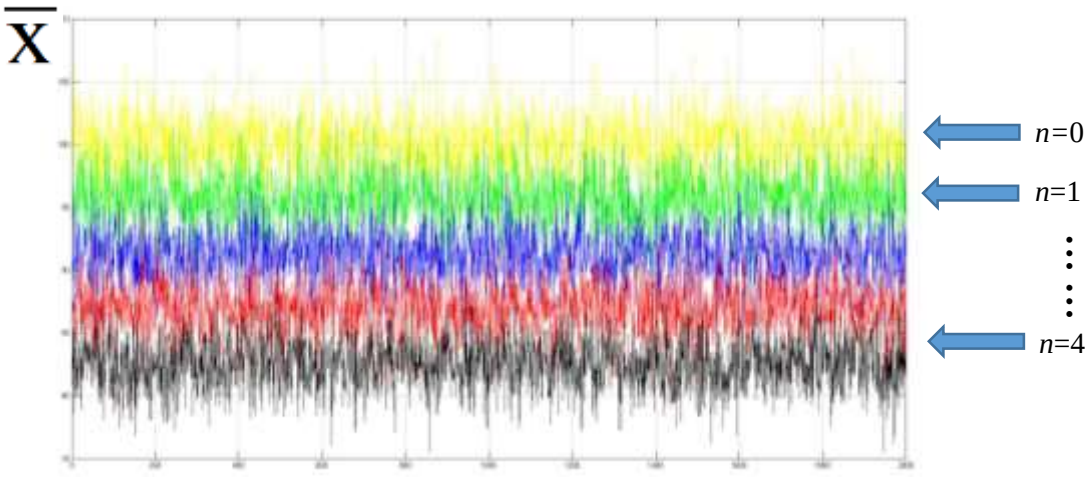
6. 计算预期“捣乱者”数量 $\boldsymbol{n}$，$\boldsymbol{n=-0.23\times\overline{x_1}+23}$

计算 n 的目的是对在首轮中因初始资金不高直接选择进行“捣乱”的那部分干扰数据进行推测，并在计算正态分布区间中最有利得分的金额时去除这部分已失去竞争力的数据的干扰.

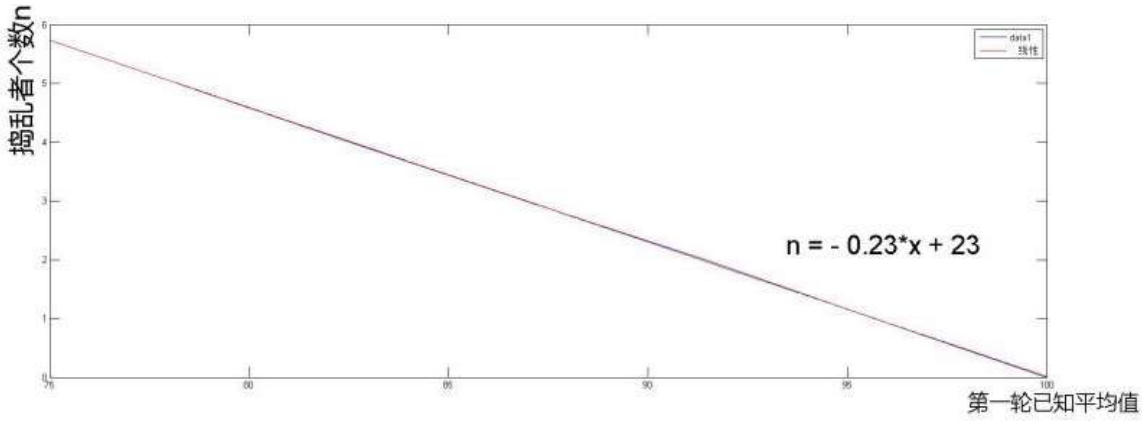
为寻找 n 与$\overline{x_1}$的关系，便于用已知值推测未知值，我们通过 Matlab 模拟来解决这一问题.

根据常识已知捣乱者个数 n 越大，第一轮已知平均值$\overline{x_1}$最小. 根据计算机模拟抽取的 2000 组样本，其分布如下图所示：

登峰杯
全国中学生学术科技创新大赛



由图验证了捣乱者数目 n 的增长在已知平均值$\overline{x_1}$的下降上有所体现.
为量化 n 与$\overline{x_1}$的关系，我们分别对 n 等于 0 至 4 时$\overline{x_1}$的值分别进行线性回归并得到五个点，n=0，$\overline{x_1}$=100，n=1，$\overline{x_1}$=95．68，…，n=4，$\overline{x_1}$=82．554．因直接将其表示为分段函数有很大可能增加段端值的误差，故我们将 n 与$\bar{x}$的函数关系拟合为连续的一次函数的形式，得到表达式$n = -0.23 \times \overline{x_1} + 23$，其图象如下：



7．计算出进入有利于得分的区间的概率 P，$P = \frac{8}{23-n}$

在得到 n 后，此时真正投标值落在本文假设的正态分布区间内的竞争者数目约为 23-n，为使得分尽量高，需要使自己排名在这 23-n 中的第 8 位附近，即处在整个最终投标值正态分布区间的$P = \frac{8}{23-n}$（从高向低数）的位置.

8．计算正态分布的标准差$\sigma = \frac{\overline{x_1}^2}{160} - 1.25 \times \overline{x_1} + 63$

我们认为，标准差与$|x_1 - 100|$呈正相关，在$\overline{x_1} = 100$时取得最小值．结合实际情况与投标价格为 0．5 的整数倍这一特点，我们认为$\sigma_{min} = 0.5$．当$\overline{x_1} = 80$ 或 $\overline{x_1} = 120$时，标准差取得最大值，我们根据经验，认为$\sigma_{max} = 3$．$|x_1 - 100|$越小，这一平均值所受干扰就越小，随着$|x_1 - 100|$的增加，其增长速度较缓

登峰杯
全国中学生学术科技创新大赛

$|x_1 - 100|$越大，这一平均值所受干扰就越大，各组的评判标准就越不一致，各组间的差距就会越来越大，这种差距的增加会越来越快．由此，我们用二次函数来模拟这一关系，代入数据后我们求出$\sigma = \frac{\overline{x_1}^2}{160} - 1．25 \times \overline{x_1} + 63$．

9．计算$1 - P$，查表得到它对应的标准正态分布的随机变量 t

由第 7 步中的论述可知，应尽量使自己排名处在整个最终投标值正态分布区间的$P = \frac{8}{23-n}$（从高向低数）的位置，即自己的两轮平均投标价 x 到负无穷区间内正态分布图象与横轴围成的面积为$(1 - P) \times 100\% \times 1$，根据标准正态分布表根据面积（概率）对应查找随机变量 t

10．根据$x_2 = 2\sigma t + 2h - x_1$求出 x_2

在步骤 9 所得的标准正态分布 t 值需返回我们所给限定条件下的两轮平均投标价x，根据$t = \frac{x-h}{\sigma}$逆推 x 的计算式为$x = \sigma t + h$．再根据 x 求出 x_2，$x_2 = 2x - x_1$．

合并两公式得$x_2 = 2\sigma t + 2h - x_1$

11．根据余额决定是否购买其他信息

首先，购买最大值、最小值和极差对我们的决策帮助不大．

由于捣乱者的存在，最小值收到大幅影响，它并不能反映有竞争力的组中的投标最小值．

我们关心自己能排到第 4—10 名，而不是第一名，因此最大值对我们用处不大．既然如此，极差对我们的用处也不大．

真正可能有用的是标准差和中位数．

标准差的用处在于，如果我们财富充足而选择购买它，那么可以估计捣乱者数量（视 9 左右为无捣乱者时的标准差，每增加 17 左右认为多了一个组捣乱），但是它未必比我们的拟合准确．

中位数的用处在于，它受捣乱者的影响较小．它可以让我们更准确地估计自己的竞争力，即首轮投标自己以上的组数．

2．5．模型的评价

2．5．1 模型的检验

登峰杯
全国中学生学术科技创新大赛

第二阶段游戏情况																	
局	财富值	第一轮	第二轮	两轮平均值	总平均值	方差	标准差	中位数	最大值	最小值	极差	h	捣乱人数	除捣乱外前八比例	二轮标准差	查表t值	前八x值	得分
1	85	0	0	0	93.22	852	29.195	96	120	0	120	94.4	1.56	0.63	0.79	0.32	94.6	0
2	109	109	95	102	101.6	0					0	101.5	-0.37	0.66	0.52	0.41	101.8	2
3	110	110	95	102.5	102.3	0		102			0	102.1	-0.52	0.66	0.53	0.41	102.4	2
4	117	117	86	101.5	94.83	783	27.985	101				95.5	1.19	0.63	0.67	0.43	95.8	6
5	120	120	91	105.5	99.87	0		108				99.9	0.03	0.65	0.50		99.9	10

由上表可以看出：

1. 在假设得到满足（即初始财富值分布均匀）时，我们通过以上方法所求出的投标价格能够让我们的平均投标价处在第八名附近．我们按照计算出的 x 值就近投标，名次在 9，10 名左右．
2. 初始财富值分布不均匀时（例如第 5 局），我们的模型不太适用，对此我们未做研究．我们对这一特殊情况进行了少许分析，较为幸运的拿到了 10 分．

由前四句的结果来看，我们希望处在第八名其实要冒很大风险；因此遇到这种情况时我们可以以投在第六或七名左右的标准来计算 x 值．

2．5．2 模型的优缺点

⑴针对此问题的模型有如下优点：

①较为简洁，易于计算；

②考虑到了多种因素，反映出实际生活中的情况；

③运用计算机模拟的方法来对捣乱者造成的影响进行量化分析，较准确、便捷的计算出捣乱者的影响；

④运用最大熵原理对第二轮的投标分布进行了分析，较为创新．

⑵针对此问题的模型有如下缺点：

①仅仅考虑到了平均值与方差这两个约束条件；

②在计算数学期望 μ 时未考虑其内在机理，仅仅根据其极值和变化率的大致趋势来得它的表达式；

③过分追求处在第 8 名获得 10 分，得分较不稳定．

3. 参考文献

[1]姜启源，谢金星，叶俊．数学模型．北京：高等教育出版社[M]，2011．1，373

[2]吴国富，安万福，刘景海．实用数据分析方法．中国统计出版社[M]，1992．10，24-27

[3]曲英杰 孙光亮 李志敏 最大熵原理及应用[J]．青岛建筑工程学院学报,1996，17(2):94-100

[4]冯尚友 信息熵与最大熵原理[J]．水利电力科技,1995，22(3)

[5]冯尚友 熵的微观解释与信息[J]．水利电力科技,1995，22(1)

4. 附录

本问题通过一个博弈型比赛项目，具体实施一个多方参与的决策过程．比赛

登峰杯 全国中学生学术科技创新大赛

分两个阶段进行，每个阶段各设 5 局比赛，比赛规则相同．具体如下：

一、比赛规则

每局比赛设有 8 份奖品，比赛开始各队首先以抽签方式获得各自财富量(范围在[80,120]内)(每个队务请注意保密)．然后根据各自财富数量(即投标价格始终不能高于你手中的财富)，通过两轮秘密投标(即每个队的投标价格不公开)，由比赛裁判组根据以下规则决定奖品的归属，并给各队打分．

•规则 1 第一轮投标结束后，裁判组公布此轮投标价格的平均值．为了在第二轮投票中争取主动，每个队可以以每个数据 5 个财富的代价获得本轮投标价的方差、极差、最大值、最小值或者中位数中的任何一个(只有本队知道，请注意保密)．

•规则 2 第二轮投标结束后，按标价从高到低排序；如果投标价相同，则扣除投标价格后剩余财富多者排名靠前；如果仍然并列，则两轮投标总用时少者排名靠前．排名最靠前的 8 个队将中标获得奖品．

•规则 3:裁判组以“中标但投标价最低为最优”为原则，根据以上各队的排序结果进行打分：第 4 名至第 8 名依次得分 2,4．6．8．10，第 9 和第 10 名得 2 分；其余各队得分为 0．

（备注：分数远比奖品重要）

二、其它说明

1．每局第一轮投标必须在裁判员宣布本局比赛开始后 10 分钟内完成，否则视为弃权；第二轮投标限时 20 分钟，自裁判员公布首轮投标均值起计时，购买统计数据的订单必须前 10 分钟内递交．

2．所有初始财富及投标价格均须为正整数每局比赛结束后剩余财富归零．

3,第一阶段得分为该阶段 5 局比赛得分的平均值(满分 10 分)第二阶段得分为该阶段 5 局比赛得分平均值的 2 倍(满分 20 分)．两阶段得分的和作为比赛总评分．

4．各队须在每个阶段开始前提交技术分析报告，说明决策过程(重点之一是数学模型的建立与分析)，以便最终答辩时作为参考．

登峰杯
全国中学生学术科技创新大赛