

## 2023年美赛C题第一问框架

### 1. 开发一个模型解释报告结果的变化规律

① 明确报告结果的数量与时间的关系

② 采用分段函数，使用线性拟合模型  
目的：得到具体的关系式

③ 验证模型准确性，即进行误差分析，作图

④ 对2023年3月1日的报告结果数进行预测

### 2. 单词的属性定义

① 具体定义，按照单词的类型定义即可

② 按单词的属性对数据分类

③ 引入统计模型，统计不同单词类型下的频率直方图

④ 描述结论即可

[以上内容不需要体现在论文中，仅用于论文的书写布局]



## C 题 第一问

### 问题分析

[为避免重复,大家需在问题分析主要体现以下点,具体听视频讲解,自行组织语言即可]

① 为了探究报告结果数量与时间的关系,可以对其进行作图描述

② 以时间作为自变量  $x$ , 以结果数量为因变量  $y$ , 对其进行散点拟合

③ 通过验证模型误差并对模型求解结果

④ 对 2023 年 3 月 1 日结果进行预测

⑤ 按照单词属性对附件中的数据进行分类

⑥ 使用数理统计方法分别统计不同属性单词不同成绩的分布,从而得到单词属性对玩家总分的影响

[有能力者可根据上述分析做解题流程图]

流程图①

流程图②

### 模型建立与解

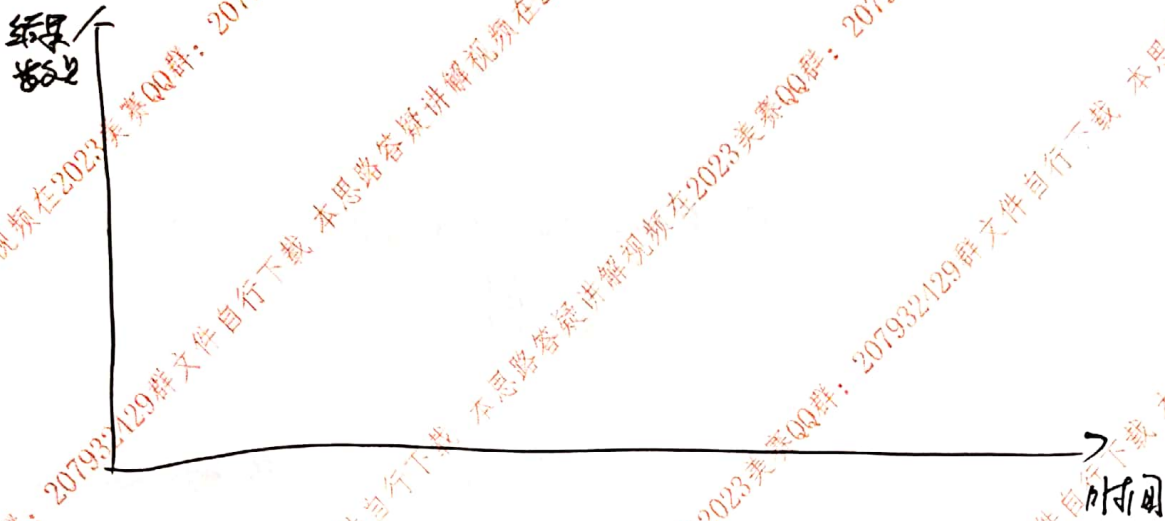
[用两三句话引出题目下文] 可解绍游戏背景

结合上述问题一的分析可知,要想探索每天报告结果数量与时间的关系,即:则需对附件 1 所提供的数据综合分析





首先做散点图如下.



由上图可以发现在随着时间的发展参与该项游戏的人数首先由少逐渐增多, 当数量达到一定高度时, 人数渐降低. 为了客观描述以上情况, 可利用分段函数进行量化, 即整个过程可分为两个阶段.

### ① 增长期

令参与该项游戏人数为  $y$ , 时间计为  $x$ . 以便于计算 取当 2022-1-21 时  $x=1$ , 当 2022-1-22 时  $x=2$ , 以此类推.

由上图观察发现该散点图符合指数变化, 假设

$$y = ae^{bx} \quad (1)$$

则通过 matlab 软件将附件 1 的数据带入对 (1) 式拟合得到

结果为  $a = \underline{\hspace{2cm}}$   $b = \underline{\hspace{2cm}}$

③



将上述结果代入原指数数据模型，存在

$$y' = ae^{bx}$$

通过计算相对误差  $w$  大小判断假设的准确性

$$w = \frac{|y' - y|}{y}$$

将附件 1 数据代入结果如下表所示：

时间	$x$	$y$	$y'$	$w$
	1			
	2			
	3			

综上所述上表可知 整个预测模型的误差较小，故此  $y = ae^{bx}$  可用描述增长期时间与人数之间的关系。

同理对减缓期也可采用以上模型。

## ② 减缓期

[格式与上述格式相同]

④





综上可得到分段函数关系:

$$y = \begin{cases} ae^{bx} & 0 < x \leq x_1 \\ a_1x + b_1 & x_1 < x \end{cases}$$

通过上述分析可得报告数量随时间的规律然而不同时期用户得分高低也会影响整个现象的数量其主要因素为单词的难度程度. 本文通过单词的属性来判断单词对用户得分的影响情况. 具体操作如下:

① 对附件中所有的单词按照属性分类.

对于单词而言. 在日常生活中常用的词大都分为名字. 一些形象词... [听视频自行组织语言描述]

② 统计每个属性下得分的人数

令某属性下单词个数为  $n$ , 其中不同单词对应的得分结果数为  $m$ .  
分别为 1, 2, 3, 4, 5, 6,  $x$  的占比为  $u_{m,1} u_{m,2} u_{m,3} \dots u_{m,7}$

则总结果数为  $M$

$$M = \sum_{i=1}^{i=n} m_i$$

得分为 1 的人数为  $P_1$

$$P_1 = \sum_{i=1}^{i=n} m_i \cdot u_{m,1}$$



得分 2 的人数为  $P_2$

$$P_2 = \sum_{i=1}^{i=n} m_i \cdot u_{m_i, 2}$$

得分  $j$  的人数为  $P_j$

$$P_j = \sum_{i=1}^{i=n} m_i \cdot u_{m_i, j}$$

则得分  $j$  的频率  $Q_j$

$$Q_j = \frac{P_j}{M}$$

③ 作频率直方图

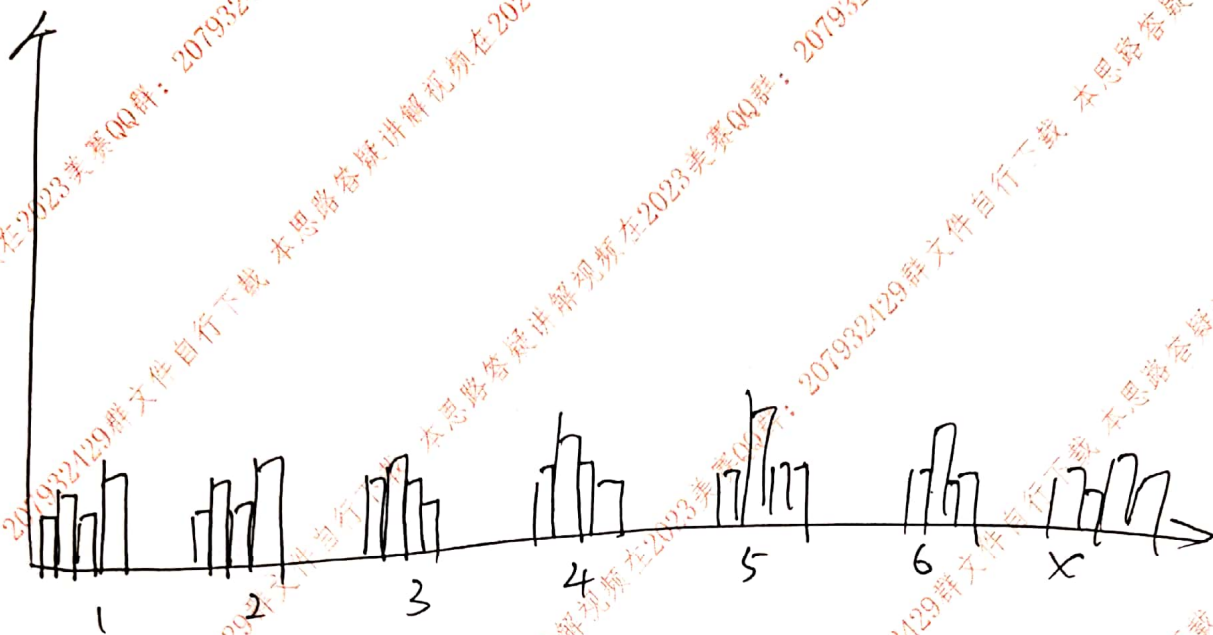
通过上表将附件中不同属性的得分频率统计如下表所示。

属性	1	2	3	4	5	6	*

通过上表作图如下:

⑤





综合描述即可：

⑦

