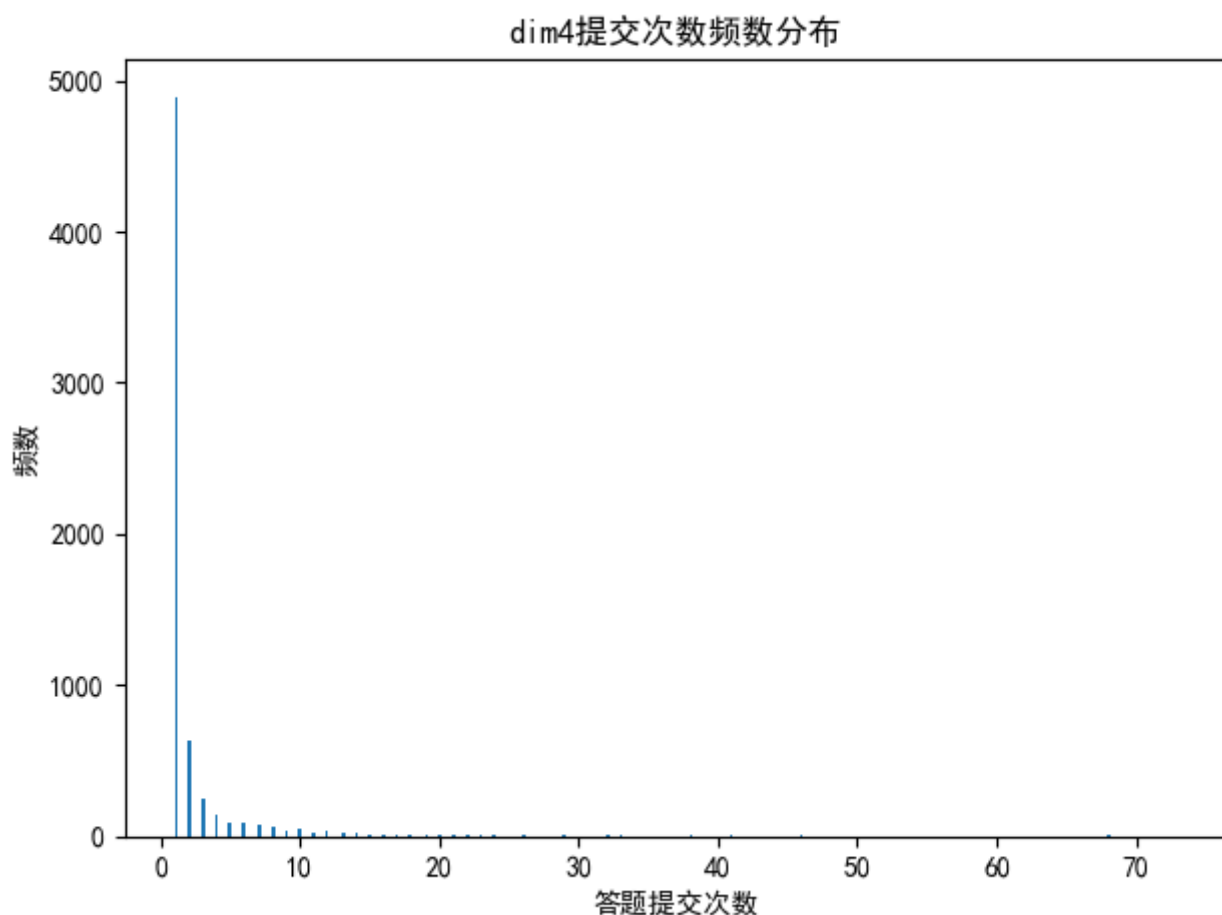


# 树、图 (dim4) 分析

## 1. 提交次数

结果阶段一的处理后，dim4有6480份样本，提交次数为1次的有4892份，提交次数为2次的有635份。



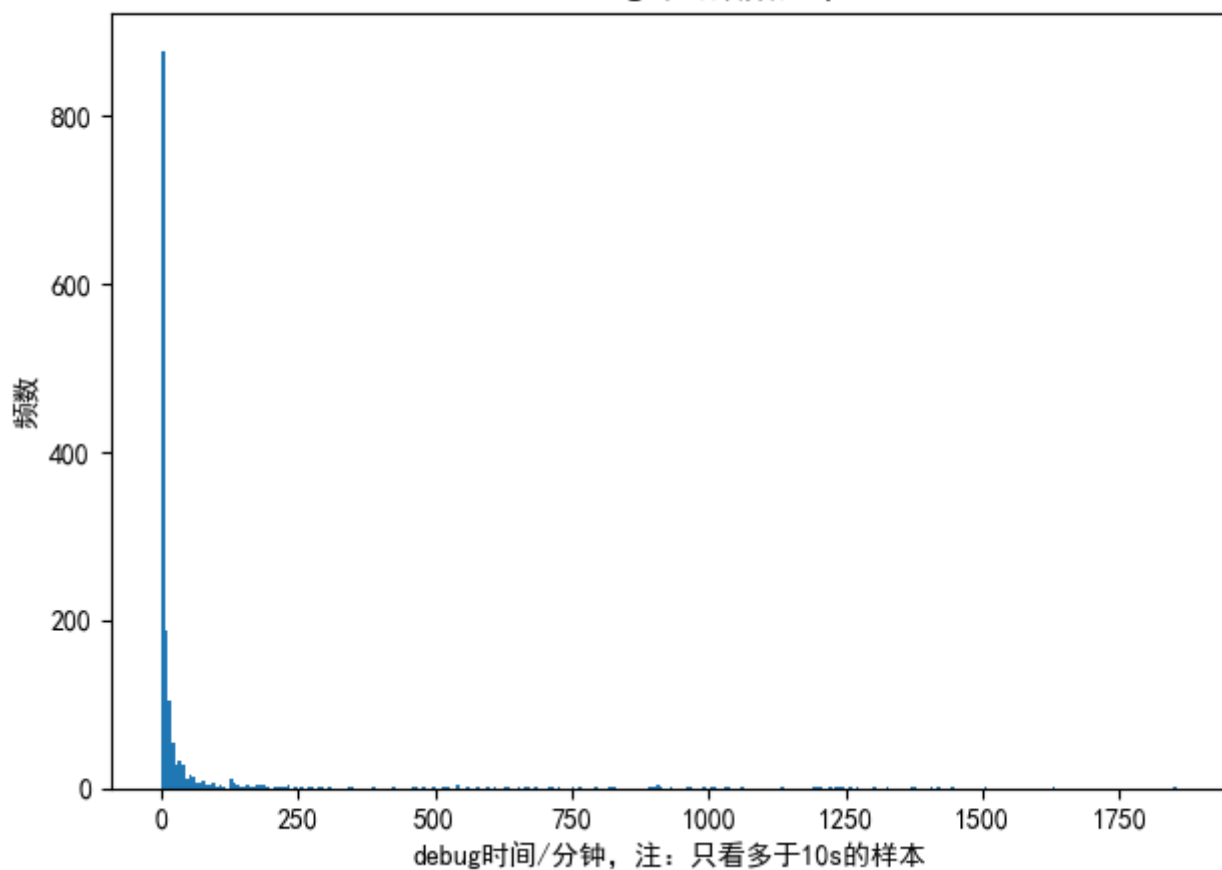
## 2. debug时间

注：最后一次提交时间-第一次提交时间，如果最后一次提交之前获得满分，则将首次获得满分的时间作为被减数。

dim4有6480份样本，debug时间少于10s的有4950份，绝大多数人首次提交便可以完成

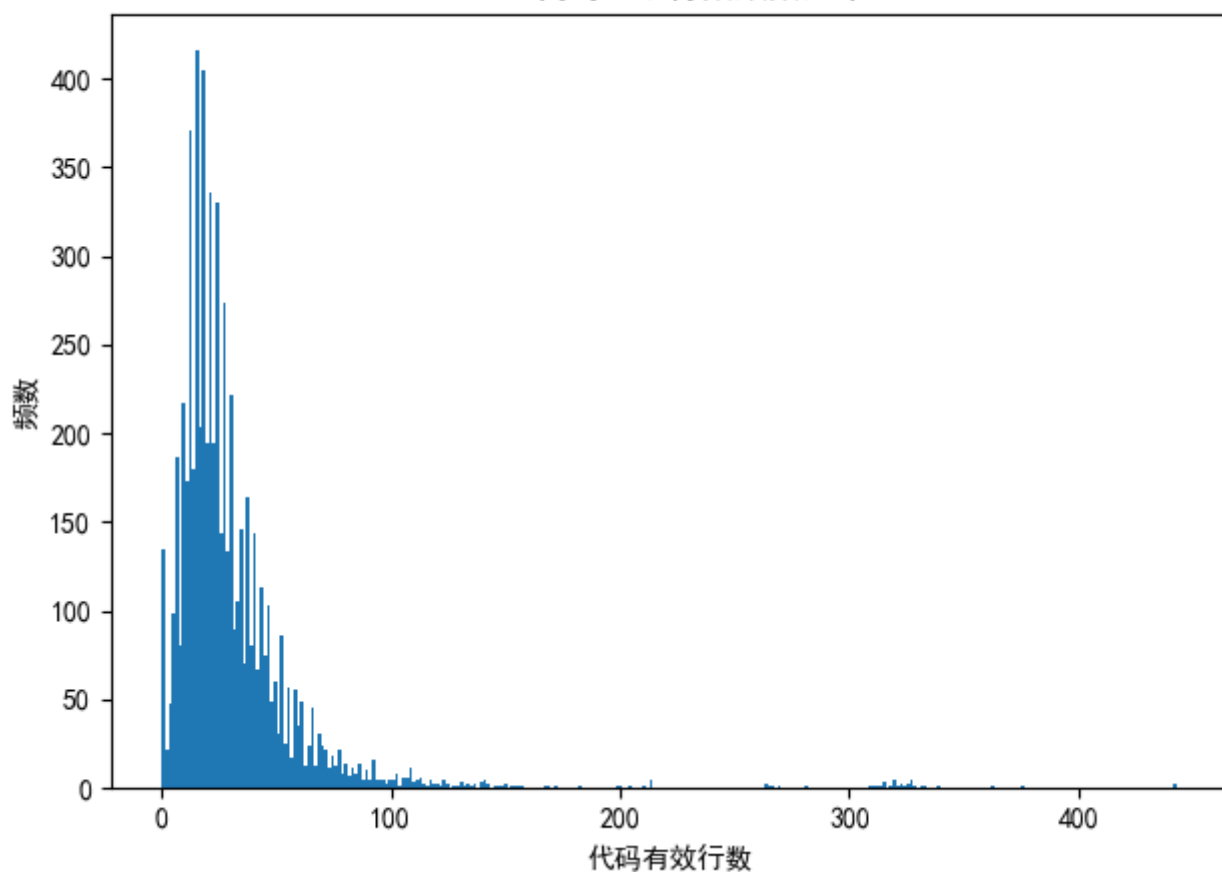
剩下的1530份样本中，大多几十分钟便可完成debug

dim4debug时间频数分布

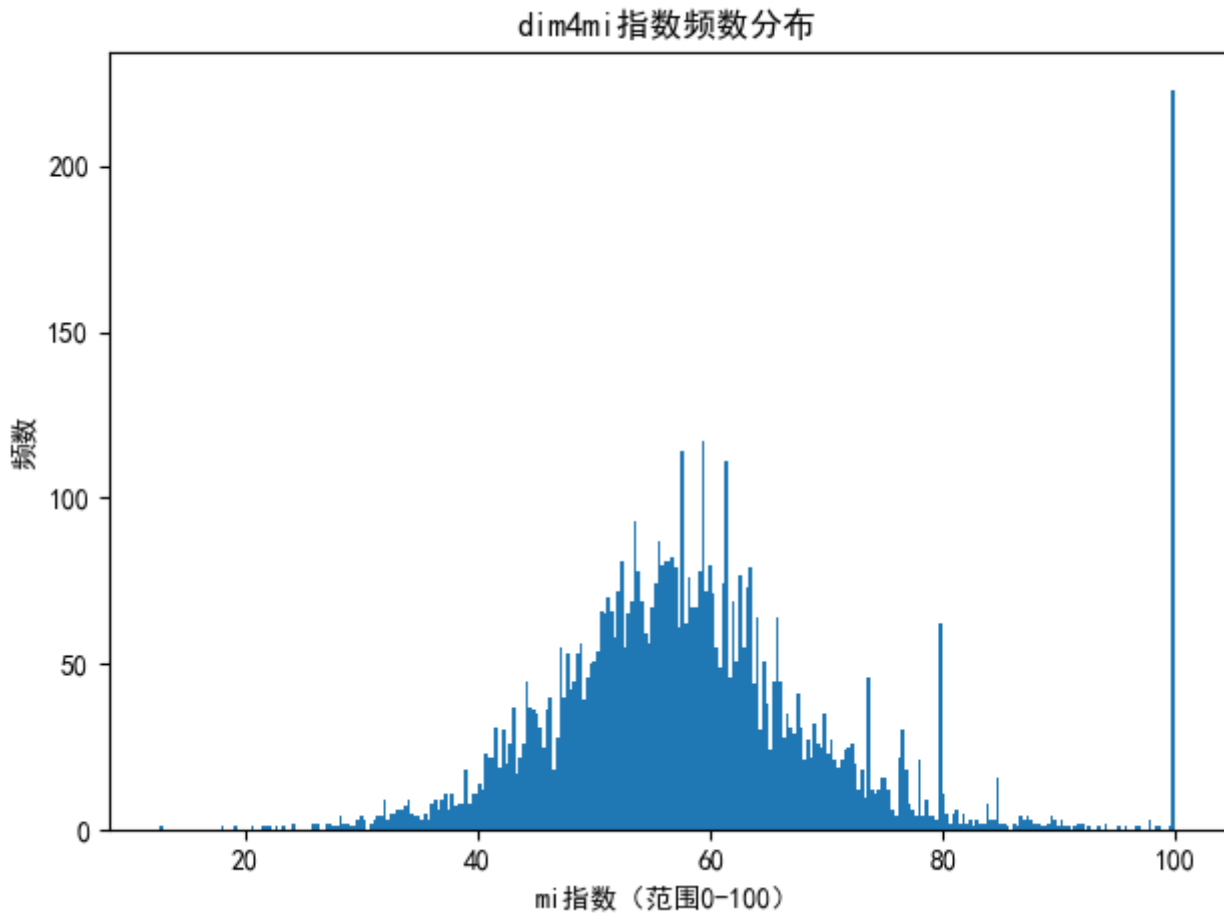


### 3. 代码有效行数

dim4代码有效行数频数分布



## 4. mi指数



## 总结

分析了提交次数和debug时间的情况，结合我们自身经验发现，大部分同学首先在本地对代码做debug，首次上传的代码往往是几乎没有bug的。

所以，一个用户的得分又各个维度得分相加而得

## debug得分

体现在提交次数和debug时间上，这里我们只考虑debug时间

我们过滤发现，debug时间小于30分钟的有6196份样本，超过总样本的95%。

用户i 做的n题中每一题的debug时间为

$$DT_i \text{ (单位: 秒)}$$

用户i 做的n题中每一题的debug得分为

$$DTS_i = \begin{cases} 1, & DT_i < 30 * 60 \\ 1 - \frac{DT_i}{30*60} * 0.01, & DT_i \geq 30 * 60 \end{cases}$$

得分为

$$DT_{dim4}^{score} = \frac{\sum_{i=1}^n DTS_i}{n}$$

**有效行数得分**

所有用户的dim4代码的有效行数均值为

$$VL_{avg}$$

用户i 做的n题中每一题的MI指数为

$$VL_i$$

得分为

$$VL_{dim4}^{score} = \frac{\sum_{i=1}^n \frac{VL_{avg}}{VL_i}}{n}$$

**mi指数得分**

所有用户的dim4代码的MI均值为

$$MI_{avg}$$

用户i 做的n题中每一题的MI指数为

$$MI_i$$

得分为

$$MI_{dim4}^{score} = \frac{\sum_{i=1}^n \left(\frac{MI_i}{MI_{avg}}\right)^2}{n}$$

**用户的树、图得分为**

$$S_{dim4}^{score} = DT_{dim4}^{score} + VL_{dim4}^{score} + MI_{dim4}^{score}$$

