6 Grade Inflation[①]

6.1 竞赛题

Background

Some college administrators are concerned about the grading at A Better Class (ABC) college. On average, the faculty at ABC have been giving out high grades (the aveage grade now given out is an A−), and it is impossible to distinguish between the good and mediocre students. The terms of a very generous scholarship only allow the top 10% of the students to be funded, so a class ranking is required.

The dean had the thought of comparing each student to the other students in each class, and using this information to build up a ranking. For example, if a student obtains an A in a class in which all students obtain an A, then this student is only "average" in this class. On the other hand, if a student obtains the only A in a class, then that student is clearly "above average". Combining information from several classes might allow students to be placed in deciles (top 10%, next 10%, etc.) across the college.

Problem

Assuming that the grades given out are (A+ . A, A−, B+, ...)

① 1998年美国大学生数学建模竞赛的B题。

can the dean's idea be made to work?

Assuming that the grades given out are only (A,B,C, ...) can the dean's idea be made to work?

Can any other schemes produce a desired ranking?

A concern is that the grade in a single class could change many student's deciles. Is this possible?

Data Sets

Teams should design data sets to test and demonstrate their algorithms. Teams should characterize data sets that limit the effectiveness of their algorithms.

赛 题 简 析

本题是学生评价问题。在美国,许多大学的管理部门很关心学生成绩的评定问题。现有的成绩评定打分过高(平均为 A-),不能区分出优秀学生和普通学生,而优秀奖学金项目仅资助 10%的优秀学生,故对学生按班进行排名是必要的。

学校校长有一种想法:在每个班上,将每个学生与其他学生比较,排出一个次序。并注意这样的因素:在一个所有学生都得 A 的班上,某个得 A 的学生只是处于平均水平,在仅有一个学生得 A 的班上,则得 A 的学生明显高于平均水平。结合各个班的信息,可以排出整个学校学生的次序[第 1 个 10%,第 2 个 10%,等等]。

在上述背景下,题目提出了四个问题:

1. 假定给出的分数为:A+, A, A-, B+,..., 校长的想法能实现吗?

2. 假定给出的分数为:A, B, C,..., 校长的想法能实现吗?

3. 其它的方法能产生希望的排序吗?

4. 单个班级的分数能改变很多学生的等级吗?

要求参赛者对上述问题做出回答,并设计数据检验他们的算法。

6.2 参赛论文

Place Students in Deciles Reasonably

Wang Guoshou Wang Yong Wu Yongjun
Advisor: He Xiaoliang

Summary

In our paper, we try to reach the goal of classifying all the students into deciles reasonably. We build up three models, which can be used in different situations.

First of all we find a method based on the psychology and the law of large numbers. With the method we discuss how many grades should be used to reach the optical effect and give a suggestion to the faculty of ABC colleges.

We use the dean's idea to design our first model named Compensation Model. In this model we discuss how the number of grades affects the result of the model.

To improve the Compensation Model, we design another model named Inflation Model. The model use information more fully than the Compensation Model and the effectiveness is improved.

To meet the need of real life, we give the third model called Precision Model to reach high precision by using some extra reasonable

information.

We design program by each model's algorithm. You can input the grades of students and the programs will give out the division of the students. These programs can be used in ABC college.

Finally we compare the three models, and characterize data sets that limit the effectiveness of each model. We test our models not only with the simulative data but with the real life data. Through the tests of the models, we discuss whether the grade in a single class could change many students' deciles in our three models.

Restatement of the Problem

Nowadays in many ABC colleges, the faculty have been giving out too high grades (the average grade given out is A−), therefore it is impossible to distinguish between the good and mediocre students. However, in many cases it is very necessary to give out a class ranking. For example, most scholarships only fund the top 10 % students.

The problem is to seek some schemes with which we can give out a ranking to meet practical needs. Certainly, the ranking must be effective and fair to everybody.

Assumptions

1. Each student is in a single class. There are several departments in a college and many classes in every department.
2. Every student has his objective achievement that can reflect his true ability. The achievement obeys the normal distribution N (75, 100).
3. Each student in an ABC college has his own everyday achievements available.
4. The importance of different achievements is not the same.
5. The grades can be quantified through some method.

6. A student's comprehensive achievement is concerned not only with his grades, but with what grades the other students get.
7. The average grade in the college is A−.
8. Our destination is to divide the students into deciles.

Notation Table

Devi	the mean square deviation of model's division. (The average is 10%).
Precision	the stable precision of model's division before combining into groups. (See the inflation model's algorithm.)

Analysis of the Problem

At ABC colleges it's impossible to distinguish all students into deciles, if we do not use extra information. The extra information must be reasonable and easy to obtain. Luckily, almost all colleges are composed of departments, and departments have several classes. According to the theory of set, if two objects are in the same set, there must be some relations between each other. Classes and departments can be regarded as two kinds of sets, therefore the students in the same class or the same department have some relations. We think of the relations as the resource of extra information. For instance, if several students are in the same class, they can compare with each other.

Because the students' achievements are given out in grades (such as A+, A, A−, B+, ...), how to quantify the grade is the first problem, no matter what approaches are used. Using the theory of psychology and the great number law, we find a perfect formula to transform the qualitative grade into the quantitative score (See the appendix: How To Quantify).

The dean's idea of comparing each student to the other students in each class is equivalent to that of comparing each student with the average of the class. Through comparing every student to the average, we can find the rank each student belongs to in a class. When combining the students in the same rank of every class across the college, compensation in each rank occurs among all the classes. When the number of classes increases, the effectiveness of the compensation can improve by the theory of the Statistics. Compensation among classes is the core of the model.

Just as the inflation, it is reasonable to suppose in a class, the bigger the number of students getting A+ is, the lower the quantified score of A+ in the class is. We assume the students' achievement obeys the distribution N (75, 100). Using this distribution and the percentage of each grade in a class, we can find the score of every grade in each class (See Figure 1). We call this score inflation score. To improve fairness and precision, we modify the inflation score with the average score in the class. In this way we can assign a score to each grade in a class, the score of A+ may be different in different class. In this way we can distinguish the same grade in different classes. So the precision is improved. We call this model Inflation Model.

Though the Inflation Model can meet the need of place students in deciles, it cannot distinguish A+ in the same class. That is to say, there exist a upper limit of precision in the Inflation Model. In some cases it's necessary to distinguish the students more precisely, for instance selecting only two students as outstanding students. So we need add extra information to reach the destination. In real life, a student's everyday grade is recorded and available, which also reflect the achievement of the student. It's reasonable to take the everyday grades into account. At the same time the everyday grades can modify the terminal grade and reflect the student's achievement more objectively. In reality this method is used comprehensively. Adopting the fuzzy set theory we can tell the difference of the same level students in the same class. Obviously the method is more precise. This is our third model named Precision Model.

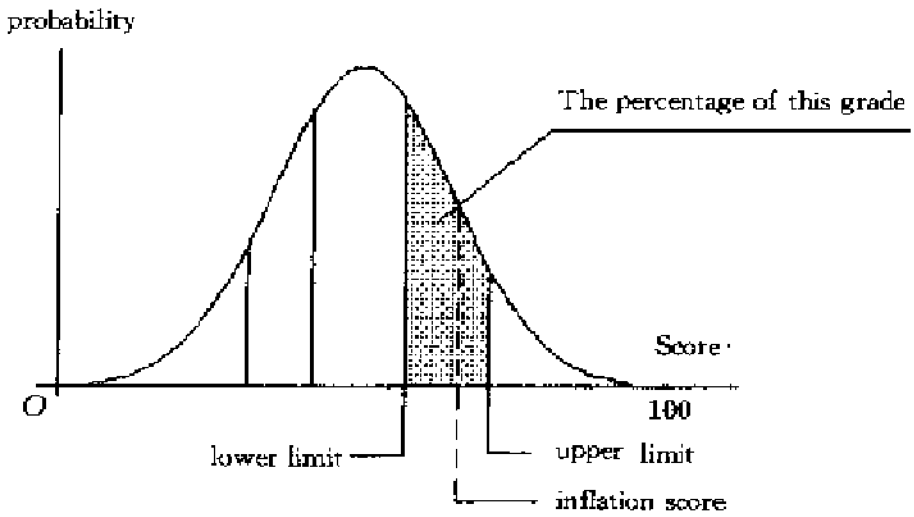

Figure 1 How to get the inflation score

According to the theory of informatics, the more information we use in a model, the higher precision the model can reach. If we can get the more comprehensive information of every student, we can reach higher precision.

Therefore the most valuable work is to seek a method that is practical and reasonable. This is the principle of our model.

Design of the Models

- The Compensation Model

In this model we design the following algorithm according to the analysis above:

1. Input the grades of a class and quantify each grade (See the ap-

pendix: How To Quantify).

2. Calculate the average score of each class.

3. Take the grade whose quantified score is closest to the average score as the average grade M (M: Middle).

4. Determine the new grade of the other score according to the average score and grade with $A1$, $A2$, ... and $B1$, $B2$, ... (A: *Above*; B: *Below*).

5. Combine the students in the same new grade of all classes into a group.

6. Output the groups in different new grades as the classification.

Some explanation of step 3, 4: (Also see Figure 2)

For instance, there are 9 grades, according to the quantified formula the quantified score set is {100, 95, 90, 85, 78, 70, 60, 48, 30}, the grade set {$A+$, A, $A-$, $B+$, B, $B-$, $C+$, C, $C-$} and the average score is 80, then the grade B is the average grade. The new grade set is {$A4$, $A3$, $A2$, $A1$, M, $B1$, $B2$, $B3$, $B4$}.

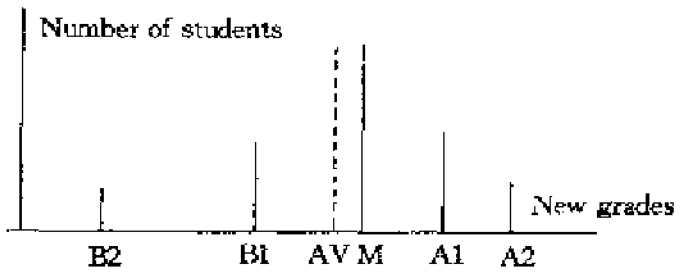

Figure 2 How to get the new grades

AV: Average

To test the compensation of the model, we change the number of classes from 10 to 20 and get table 1.

First we let the grade set be {A+, A, A−, B+, B, B−, C+, C, C−}, the number of grades is 9.

Table 1

The number of class	20	18	16	14	12	10
precision (%)	12.0	12.3	13.6	14.0	14.5	15.0
devi (%)	3.12	3.20	3.30	3.40	3.50	3.60

From the table, with the decreasing of the number of classes the precision decease slowly. But the deviation is stable

To test the stability of the model, we let the variance change from 6 to 14 and get the table 2.

Table 2

variance	14	12	10	8	6
precision (%)	11.1	11.5	12.2	12.3	13.2
devi (%)	2.86	2.96	3.03	3.19	3.33

From the table, with the decreasing of the number of classes the precision decease slowly. But the deviation is stable.

Then we let the grade set be $\{A, B, C\}$, the number of grades is 3. The following is the result of test:

Table 3

The number of class	20	18	16	14	12	10
precision (%)	20.0	21.3	22.5	24.5	25.0	25.8
devi (%)	13.6	14.1	14.5	15.0	15.4	15.9

Table 4

variance	14	12	10	8	6
precision (%)	23.8	27.4	31.4	35.2	37.4
devi (%)	12.5	13.2	13.9	15.0	15.7

Obviously when the number of grades is 9, the model can work properly. If the number of grades is 3, the model's effectiveness is bad.

Using num stands for the number of grades, let num = 8, 6, 5, 4: test the model, the result is the following (let the variance be 10, the number of class be 20).

Table 5

num	8	6	5	4
average precision(%)	13.6	15.3	16.8	18.5
average devi(%)	5.09	7.21	9.01	11.3

That is to say with the decreasing of the number of grades, the precision become worse and the deviation increase.

- **The Inflation Model**

Though the first scheme is easy to operate, its precision is not good and its effectiveness depends on the number of classes greatly. We form the second model so as to distinguish the students more precisely and effectively.

We think the achievement of students in a class obeys the normal distribution N (75, 100). So in a class the more students getting A+ are, the lower the quantified score of A + is in the class. That is to say, high grades will inflate if a lot of students get high grades in a class. According to the thought, we obtain the following algorithm:

1. Input the grades of the students in a class.
2. Calculate the percentage of the students in each grade.
3. Get the upper limit and the lower limit corresponding to every grade, according to the percentage of the grade and the distribution N (75, 100). We take the upper limit of the highest grade as 105

and the lower limit of the last grade as 45 (*See figure* 1).

4. Take the average of the upper limit and the lower one as the score of the grade in the class (See figure 1).

5. Process the grades in other classes in the same way as Steps 2, 3, 4.

6. Sort the scores in increasing order across the college.

7. Combine the students getting scores similar to each other into a group, with the principle that the percentages of the group must be close to 10% to the best degree.

8. Output all the groups as the result of the model.

9. The following is the test result:

We test the model by letting the number of classes change from 10 to 20 and get the table 6.

Table 6

The number of class	20	18	16	14	12	10
Precision (%)	7.49	7.65	7.87	8.13	8.65	8.79
devi (%)	3.31	3.48	3.55	3.67	3.87	3.99

To test the stability of the model, we let the variance change from 6 to 14 and get the table 7:

Table 7

Variance	14	12	10	8	6
Precision (%)	6.41	7.53	8.31	10.4	15.1
devi (%)	3.18	3.45	3.33	3.27	3.19

Comparing the table 6, table 7 with table 1, table 2, we can find Inflation Model is more precise and effective than Compensation Model.

From the result of test we find in some cases the students getting the same grade in different classes cannot be distinguished. We try to avoid the case in the following improvement.

Improvement

1. Take a department as a class and get the scores of the grades as reference scores with the algorithm above.
2. Weighted average the score in a class and the reference score as the final score
3. Process the final scores as the step 6, 7 in the algorithm above.
4. Output the result.

The following is the test result:

Table 8

The number of class	20	18	16	14	12	10
precision (%)	7.30	7.52	7.70	7.91	8.15	8.39
devi (%)	2.31	2.43	2.57	2.71	2.89	3.02

That is to say, with the decreasing of the number of class, the precision and deviation change slightly.

To test the stability of the model, we let the variance change from 6 to 14 and get the table 9.

Table 9

variance	14	12	10	8	6
precision (%)	6.12	7.24	8.00	9.16	11.2
devi (%)	2.21	2.33	2.41	2.71	2.96

That is to say, the precision decrease quickly with the change of variance, the deviation increase slowly.

Comparing the table 8, table 9 with table 6, table 7, we can find the improvement is more precise and effective than original Inflation Model

Some Suggestion

The algorithm can distinguish the students precisely, but the fairness is not good, hence some students may complain that their classmates is so excellent that they can't get a scholarship though they are also very good students. In order to get rid of the complaint to some extent, we can modify the scores of the grades in different classes according to the average quantified scores of the classes when we use the algorithm in reality.

- **The Precision Model**

According to real research in our college, the everyday grades of students are considered in some degree when the faculty rank students every semester. So it is reasonable to take the everyday information into account.

We adopt the fuzzy set theory to process this information.

Symbol notation:

$L = \{l_1, l_2, l_3, \ldots, l_n\}$ is lesson set, l_i is the lesson i that a student study in a term, $i = 1, 2, 3, \ldots, n$.

$R = \{r_1, r_2, r_3, \ldots, r_m\}$ is rank set, r_j is the ordinal number of each rank, $j = 1, 2, 3, \ldots, m$.

$T = \{t_1, t_2\}$ is test set, t_1 means the everyday tests, t_2 stands for the terminal tests.

The volume of L is changeable according to the actual situation. The precision needed determine the volume of R.

According to the theory of fuzzy set, to tell which rank every student belongs to we must obtain a weight vector, then classify the student to a rank using the principle of fuzzy. The following is the algorithms.

1. Input the everyday grades and terminal grades.

2. With the everyday grades of a student, we create the original matrix $R^{(1)} = (r_{ij}^{(1)}) \cdot r_{ij}^{(1)} = n_j / total$, n_j is the times of everyday tests in which the student rank of lesson l_i is r_j, total is the number

of everyday tests. $R^{(2)} = (r_{ij}^{(2)})$,

$$r_{ij}^{(2)} = \begin{cases} 1, \text{ if the rank of lesson } l_i \text{ is } r_j \\ 0, \text{ if the rank of lesson } l_i \text{ is not } r_j \end{cases}.$$

3. Determine the weight vector of tests $W_t = (w_{t1}, w_{t2})$ and the weight vector of lessons $W_l = (w_{l1}, w_{l2}, \ldots, w_{ln})$. Empirically, $W_t = (0.6, 0.4)$. The vectors W_l can be gotten from the actual case, such for the specialization of the student.

4. Let $r_{ij} = \sum_{k=1}^{2} w_{tk} * r_{ij}^{(k)}$, $R = (r_{ij})$ be the judge matrix.

5. Let $S = (s_1, s_2, s_3, \ldots, s_m)$ represent the fuzzy comprehensive evaluation vector, $S = W_l R$.

According to the principle of *the fuzzy set theory*, if s_k is the maximum in set $\{s_k\}$, the student belongs to r_k.

6. Modify the evaluate vector. Let s_k stands for the maximum of $\{s_i\}$, $i = 1, 2, 3, \ldots, m$.

(a) If $\sum_{k=1}^{k_0-1} s_k \geqslant \alpha \sum_{k=1}^{r} s_k > \sum_{k=k_0+1}^{r} s_k$, the rank fall by one;

(b) If $\sum_{k=k_0+1}^{r} s_k \geqslant \alpha \sum_{k=1}^{r} s_k > \sum_{k=1}^{k_0-1} s_k$, the rank rise by one;

Empirically, $\alpha = 1/2$.

7. If there are several equal maximums, we choose the highest rank as the one of the student approximately.

8. Output the rank that the student should belong to.

We design a program to realize the algorithms. To test the effectiveness of the model we collect the real records of students in our college, the result is in accordance with the real rank very well.

The use of the model is flexible. We can use the model independently if the data input are the students_ terminal and everyday grades. In most cases, it can be used to distinguish the students getting the same grade in the same class after the processing of the second algorithm. The model uses a student's almost all recent information available, so its reliability is high. We can use it in reality.

Some Discussion of the Models

We test our models in two aspects. First we test them in a computer; then we get some real life data to test the models in reality.

- **How Many Grades the Faculty should Use in Reality**

The average grade now given out at ABC college is an A−. According to the distribution $N(75, 100)$ of the students' achievements, we can know the achievements is almost all in the interval [45, 105] and the average grade stands for 75. According to the Appendix How to quantify, if the faculty adopt five grades (A+, A, A−, B+, B), they stand for (100, 90, 77, 61, 39). We can conclude that almost all students are in the five grades and five grades is enough to evaluate the students in ABC college. In the same way we can know if the faculty adopt the grades (A, B, C, ...), the four grades (A, B, C, D) are enough. Certainly, if the faculty want to distinguish the students more precisely, they should adopt the grades (A+, A, A−, B+, B).

- **How to Simulate in a Computer**

In a computer we use the normal distribution $N(75, 100)$ to produce the original achievements of students. Then we transform the achievements into grades according to the quantified scores of each grade. Process the grade set in the college with the first and second algorithms.

- **A Real Life Test**

In order to test Assumption 2 and the Compensation Model, we get the achievements of different classes in our college, which are given out as scores. From them, we find that the average achievements of different classes are approximately 73.8 and the variance are about 11.0, that is to say Assumption 2 is reasonable. To test our models, first we transform the scores into the grades according to the quantified score of each grade. Then we process the grades in a class with the first algorithm and find that it can reach the goal of classifying the students into deciles, while its result has some difference with the true deciles.

• **Compare the Three Models**

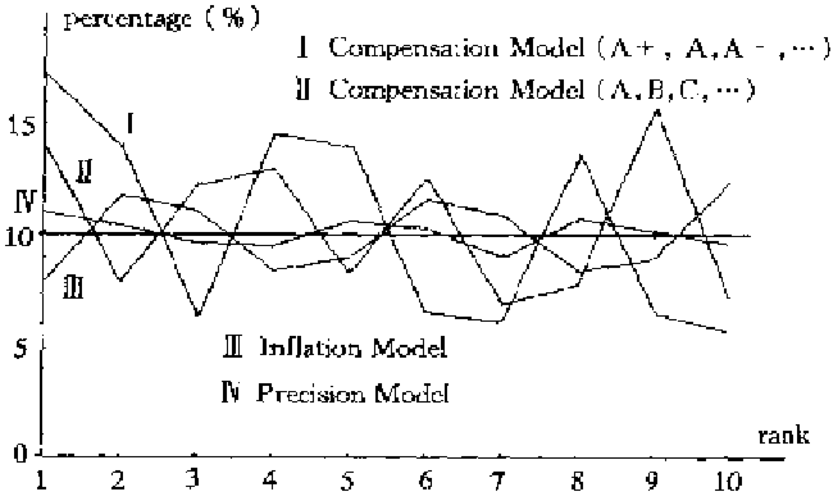

Figure 3 Comparison of models

We do simulations for different models in a computer, and record the stable percentage of each rank. Then we plot the result in figure 3. In the simulation, the mean is 75, the number of the class is 20.

• **Limitation of the Models**

The bigger the number of the classes is, the more effective the Compensation Model is. At the same time if there is little difference among the average scores of the classes, the effectiveness will also decrease. Certainly, the method is always unfair to some students.

When there is too much difference among the true average levels of the classes, the Inflation model cannot work properly (See figure 4), hence the fairness is very bad in the condition. On the other hand, if the percentages of the students getting the same grade in different classes are similar to each other, the precision and effectiveness of the model will decrease.

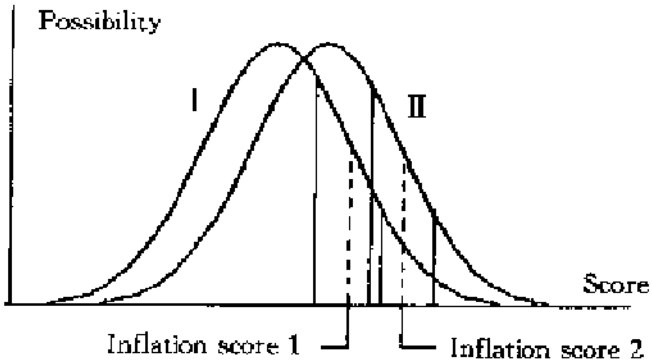

Figure 4 Wrong inflation score occur

I Get the right inflation score

II Get the wrong inflation score

The Precision Model is not suitable to place students in deciles when the number of the students is very large. But it is very effective to distinguish the students that the other model cannot divide.

Strengths and Weaknesses

The merit of the first model is simple and easy to operate, but its effectiveness depends on the number of classes in the college greatly, at the same time its fairness is not very good. In reality, if we need not distinguish the students precisely, we can adopt the simple method.

The Inflation model use the same extra information as the first one, but process it in the inflation way. It can distinguish the students more precisely. Of course, the algorithm is of more complexity. If we want to distinguish the students in different departments more reasonably, we can use the improvement model of it. After the improvement it is still unfair to a small part of students. So we give some suggestion to the dean when he uses the algorithm. The suggestion can get rid of the unfairness to some extent.

We use a student's everyday grades in the third model. The main merit of the model is that it can be used in different situations. If we want to improve the precision of the second algorithm, we can use the third model after the processing of the second algorithm. It can distinguish the students getting the same grade in the class perfectly. On the other hand, it can be used independently, while the results is not very good. In the model, we use the fuzzy set theory and the complexity of it depends on the number of the students in the same grade.

Comparing the three models with each other, we conclude that the more extra information we use in the problem, the more precisely we can distinguish the students, the more complex the algorithm is. So we should select one of the three models according to practical needs.

A concern of the problem

In the Compensation Model the grade in a single class can only affect the deciles of the students in the same class according to the algorithm, that is to say in this model, the grade in a single class could not change many students' deciles. We do a lot of simulations in a computer and the result proves our analysis.

In the Inflation model it is not so. If the grade in a class changes, the Inflation score of each grade in the class will change. Therefore, after we sort the Inflation score across the college, the orders of many students may be different with the previous ones, that is to say the grade in a single class will change many student's deciles. Assuming that there are 1000 students and 20 classes in a college, we do simulations in a computer and find that there are about 207 students whose deciles change.

Because the third model is used in most cases after the processing of the Inflation model, the grade in a single class has the same influence on the student's deciles as is in the inflation model. If it is used independently, the new grade of students only reflect the students ability more objectively, after the processing, so there is not the concern in the model.

From the analysis above we find whether the grade in a single class could change many students' deciles depends on the methods. There is not a definite answer to the question. We can take it as a factor to evaluate the models.

Our Suggestion

Using the first model, we can only place the students in deciles approximately in the inflation model, we can distinguish the students with more precision, such as in 7%. After the processing of the second algorithms, if we use the precision model, we can divide the students into smaller groups.

Appendix

- **How to Quantify**

In order to get as more information as possible, it's natural to quantify the grade of every student, no matter what approaches we use. The common method of quantification is to assign a determined score for each grade. The following is an example:

Table A1

Grade	A+	A	A−	B+	B	B−	C+	C	C−
Quantified Value	95	90	85	80	75	70	65	60	55

Obviously, this method of quantification is not ideal enough to meet practical needs.

How to quantify the grades is a very important step in this problem, we get a perfect formula according to fuzzy set theory, the law of large numbers, and Psychology.

We assume that the change of a person's sense to an object or a system is direct ratio to the change of the object or the system.

Let S represent the sense of the person, T stand for the situation of the object or the system. The change rate of the objective situation is $\Delta T/T$. According to the assumption above we can get:

$$\Delta S = K * \Delta T/T$$

Integrate the equation:

$$S = K * \ln T \tag{1}$$

That is to say, the sense of a person is logarithm ratio to T.

Through the processing of people's brain, qualitative rank (I) is available such as A+, A, A−, B+, B, B−, ... Let I describe the change of the objective situation, assuming that:

$$T = f(I) = X \pm I$$

X is an undetermined variable, according to (1):

$$S(I) = K * \ln f(T) = K * \ln(X \pm I) \tag{2}$$

Assuming that S is in the interval $[0, 1]$, $I = 1, 2, 3, \ldots, j, j+1$.

Using the two boundary conditions:

1. $I=1$, $S=1$; $K * \ln(X \pm 1)=1$ (3)

2. $I=j+1, S=0$: $K * \ln[X \pm (j+1)]=0$ (4)

Solve the equations (3), (4):

$X=j+2$,

$K=1/\ln(X-1)=1/\ln[(j+2)-1]$ (5)

when the sign in the equations is + , the solution have no real meaning.

Let $m=j+2$ and substitute (5) in (2), we can get the function:

$$S(I) = \ln(X - I)/ln(X - 1) = \ln(m - I)/ln(m - 1) \quad (6)$$

In the function, $i=1, 2, 3, \ldots, j, j+1$, $S(I)$ is the value of the function corresponding to i, j stands for the number of grades, m is a common parameter.

The function (6) expresses the relation between the definite quantitative value and qualitative value, its base is concrete and we think the result of the method is reasonable.

Table 2 number of rank = 9

I	1	2	3	4	5	6	7	8	9
$S(i)$	100	95	90	85	78	70	60	48	30

Table 3 number of rank = 5

I	1	2	3	4	5
$S(i)$	100	90	77	61	39

Table 4 number of rank = 3

I	1	2	3
S(i)	100	68	43

The above are four examples to demonstrate the methods. In some situations to take the $S(1)$ as 100 % is not suitable, suppose the highest $S(i)$ is max, max$\in(0,1)$, we may substitute (3) with

$$I = 1,\ S = \max:\ K * \ln(X \pm 1) = \max \qquad (7)$$

Thus we can also get a function:

$$S(i) = (\max * \ln(x - i)) / \ln(x - \max) \qquad (8)$$

For instance, let max=95, we can get:

$$S(i) = (0.95 * \ln(x - i)) / \ln(x - 0.95) \qquad (9)$$

With equation (9), we can obtain the following tables:

Table 6 number of rank = 9

I	1	2	3	4	5	6	7	8	9
$S(i)$	95	90	90	86	78	74	60	57	29

Table 7 number of rank = 5

I	1	2	3	4	5
$S(i)$	95	85	73	58	37

Table 8 number of rank = 3

I	1	2	3
$S(i)$	87	69	44

As shown above, if the highest $S(i)$ is not 100 % , when the number of the rank decrease to a small number, the real highest $S(i)$ is less than the given max.

According to the analysis of the above, this quantified method is perfect when the highest quantified value is 100 %, but if not so, this approach can work without any problem only when the number of the ranks is high enough.

论文点评

本文作者建立了三种模型：

1. 首先，按照一种合理的等级量化标准，提出了一个比较简单的划分模型。这个模型采用线性的划分方法，通过求线性平均的方法得到中值，然后以中值为参考分别得到各个分数段的成绩。这个模型的特点是操作方便，基本上可以满足问题的要求。

2. 其次，在第一个模型的基础上，针对其缺点，用概率统计的方法对数据进行校正（所谓的膨胀，比如一个班级中 A 等级越多，该班级 A 等级越贬值——这个假设有一定的合理性）。这个方法可以避免第一个模型中一些不公平现象的发生，使不同的班级的成绩具有可比性。这个模型的特点是可以在比较复杂的学生群体之间进行比较。

3. 最后，在第二个模型的基础上，进一步地使用模糊数学的原理和方法，把学生平时的学习情况加入到成绩的评定中，在多维平面上进行综合聚类。因为平时的成绩可以比较稳定地反映一个学生的长期学习表现，所以可以使整个模型更加可靠和稳定。

本文的特点是由粗到精，逐步深入，可以根据不同的实际情况选择不同的模型以得到比较好的应用效果。本文所建模型合理，论证正确，表述清晰。论文中设计的方法在结果上还不够好，对学生的分类还不够公平，这是本文的缺陷。

本篇论文获得 1998 年美国数学建模竞赛二等奖。