# A Mathematical Programming Approach To Aggregating Raw Scores Into A Course Grade

**Badreddine Jerbi**

**Abstract:** Teachers use tests, quizzes, assignments, labs, exams… to evaluate the students' learning. This assessment leads to grades which are vital to students, their parents and teachers. The ways the scores are combined are diverse and differ from one course to another and different ways of combination may result in different grade. What is precarious is to see that a method of aggregation can result in a failure of a student and another method produces a success for the same student. In this paper and based on the results in a semester course, I try to find the best weights that have to be allocated to the tests, the quizzes, assignments, labs and exams with a a mixed linear program that if were used, they would have maximized the number of successes. An immediate benefit is to implement those weights for the following semester the course is offered.

**Index terms:** learning assessment, grade system, scores, aggregation, mixed integer program

## 1. INTRODUCTION

[1] reported that the Alliance for Excellent Education estimated the annual cost of high school failures in the United States exceeds $330 billion and the way to solve this problem is to implement an effective grading system. This not only results in lower failure rate but also in an improved student behavior and a better faculty morale as well as an increase in teaching advanced courses. Usually teachers follow the system in place at an educational establishment to give students grades when they begin working. Most of the time the allocation of each weight to a homework assignment, a midterm, or an exam is inherited from the previous semesters syllabi of the course. Meeting of the teachers of a particular course can discuss how to motivate students by reassignment of the possible points of each demanded work. This results in a new syllabus which states the new contract between the teacher and the students on the necessary number of points to succeed. And even by doing so, there will be failing students .In this paper, I deal with this problem. In section 2, I give few recommended methods in the literature to aggregate scores into grades, in section 3, we state our problem and what are the types of students to consider, in section 4, I give the methodology to solve the problem by illustrating it on an example. In section 5, I give the results and discuss them and suggest an improvement of the solution.

## 2. LITERATURE REVIEW

For students, grades are very important, for teachers the way grades are allocated to motivate the students is a source of anxiety because it takes time and a lot of thinking. The grading begins with testing, the teacher must ask himself or herself on the reasons for and the frequency of testing, what to be covered on the tests, how to write test problems and types of questions: short-answer, long-answer, multiple-choice, What are the necessary steps to make the students comfortable during the administration of the exam. Another component is the scoring part; how to assign scores to each question and how fair that is. Finally grading the assignment or the exam; whether to assign partial credits for example [2]. Professors use a variety of assessments, such as tests, quizzes, homework, projects and oral presentations to evaluate student mastery of subject matter. Different methods of aggregation can result in different course grade. There are three famous aggregation methods: Total Possible Points" supposes that there are a certain number of possible points to be collected in the whole course. Each assignment or test has an allocated maximum number of points, with the instructor assigning more points to those items considered more important. This is mathematically similar to taking a weighted average of numerical scores of the assignments. The second method is the Weighted Average Letter Grade method. It converts the score on each assignment or test into the numerical equivalent of a letter grade, using the 4.0 scale. The instructor then calculates a weighted average of the various assignments and tests. This method of combining assignments within a course is consistent with how schools combine grades across individual courses into overall student grade point averages but does include a loss of information as compared to the preceding method. The third method is the Median method of aggregating assignment grades where different assignments have different weights, an instructor using this process would adjust the process to count the assignments with higher weights more heavily and take the median [3]. In medical education grades are often determined by combining the scores of different assessments. While the psychometric literature recommends that the variance in each assessment be identical, there is limited discussion of the impact of different methods of combining scores. Various methods of combining scores were then used. There is a large variation in the number of students failing associated with different methods of combining scores. Choosing a method to produce a combined score is a non-trivial exercise. The impact on student progression can be significant and whatever method is chosen will require justification see [4]. The term marks can have the meaning of the concepts of scores (raw test performance) and grades (level of performance). [5] considered different methods of aggregation of scores into a grade. They suggested the following steps: Scores should always be converted to grades before aggregation. The process of converting scores to grades requires that the scorer must be skilled in the subject and familiar with the principles of education. The grade scale used should be different from scores. The overall performance is computed on the basis of the median and not the arithmetic mean. The interquartile range is

- *Badreddine Jerbi is currently Associate professor at the Higher Institute of management in Gabes, Tunisia detached to the college of Business & Economics at Qassim University, Saudi Arabia, 52571. E-mail: B.jerbi@qu.edu.sa*

taken as indicator of measure of variation. After each assessment, the student has to know his/her score and grade, eliminating the case of aggregation of many assessments score. Where possible, assessment should report performance by individual assessment, not by a single aggregated mark. Students not performing well should have particular attention during assessment. Important decisions about students are made by combining multiple measures using different decision rules. Many researches concentrated on methods to compare the accuracy of single measure decision schemes. [6] argues that such methods are not suiTABLE for estimating the accuracy of decisions based on multiple measures. They presented a simulation method useful for estimating classification accuracy for any measurement model and complex decision rule. The method is illustrated by application to actual data from General Educational Development (GED) test takers. Findings in this study show how multiple scores that are combined can have a large effect on the accuracy of the resulting decision.

In their paper, [7] consider an important topic of fairness in grading. They represented each grade as a student's native performance minus the course's difficulty plus a statistical error. They consider two statistical methods for assigning an aptitude to each student and, simultaneously, a measure of difficulty to each course: (1) they minimize the sum of squares of the errors, and (2) they minimize the sum of the absolute values of those errors. They argue that by accounting for course difficulty, they arrive at a measure of aptitude that is fairer than the usual grade point average metric. Simultaneously, the measures of course difficulty can be used to inform instructors as to how their courses compare to others.

Operational Research (OR) techniques have been used, for a long time since the appearance of the discipline, to many problems in education. At the government level, OR have been used to the allocation of resources to different sectors and institutions. Another issue concerns the efficient operation of institutions, how to measure it, and how to encourage efficiency savings. OR have also been used to solve problems for the creation and location of new institutions or closure of existing ones, as well as the daily logistics of transporting students to schools. Issues of concern for school managers involve budgets' allocation, Timetabling. For the problems and the OR techniques used to solve them see the surveys of [8] and [9]. Although these reviews have addressed many of the problems and techniques, almost nothing is said about the vital problem of learning assessment. We now develop our own scheme of grading.

## 3. PROBLEM STATEMENT

The student's grade might be sensitive to the possible points given to each exam/assignment. It is the purpose of this paper to consider all the scores of the students of the course and try post adjust the possible points for the students and to find the weights for each exam/assignment those if were used they would have given more successes. These weights could be used for the following semester the course is given in order to obtain the best distribution of points that give most successes.

## 4. METHODOLOGY

### 4.1 Preface

Usually the teachers allocate the grade points for what they think it is appropriate to measure the learning abilities of the students; for example, if I take a course that has a first midterm, a second midterm, quizzes, lab and a final exam (see TABLE I).

**TABLE I : Initial points' distribution**

| Exams | First midterm | Second Midterm | Quizzes | Lab | Final Exam |
|---|---|---|---|---|---|
| Possible points | 15 | 20 | 10 | 5 | 50 |

May be there is a better distribution suiTABLE for students: Suppose I have 8 students and they are reported to have the following scores expressed as percent (see TABLE II). If I calculate the number of points gathered by the first student for example, then the student obtains the result of 0.6(15) + 0.65(20) + 0.7(10) + 0.8(5) + 0.8(50) =73

**TABLE II: Percent of points realized by each student on each exam**

| Exams | First midterm | Second Midterm | Quizzes | Lab | Final Exam |
|---|---|---|---|---|---|
| 1 | 0.6 | 0.65 | 0.70 | 0.80 | 0.8 |
| 2 | 0.47 | 0.50 | 0.8 | 0.80 | 0.66 |
| 3 | 0.60 | 0.50 | 0.60 | 0.40 | 0.64 |
| 4 | 0.50 | 0.40 | 0.50 | 0.40 | 0.50 |
| 5 | 0.65 | 0.30 | 0.60 | 0.50 | 0.48 |
| 6 | 0.45 | 0.35 | 0.46 | 0.37 | 0.5 |
| 7 | 0.50 | 0.60 | 0.40 | 0.70 | 0.45 |
| 8 | 0.60 | 0.60 | 0.70 | 0.70 | 0.6 |

And hence the number of points accumulated by each student is given in TABLE III.

**TABLE III: Cumulative points of each student**

| Student | Points accumulated |
|---|---|
| 1 | 73 |
| 2 | 62.05 |
| 3 | 59 |
| 4 | 47.5 |
| 5 | 48.25 |
| 6 | 45.2 |
| 7 | 49.5 |
| 8 | 61.5 |

Assuming that the number of points to succeed is no less than 60, we notice that the student 1, 2 and 8 have succeeded. In our example and after accepting all the students who made a score more or equal to 60, we reconsider the failing students (who made less than 60). An alternative distribution of exam/assignment points could give better results, and a more realistic way would be to vary the possible points of each exam/assignment on an interval rather than considering one point. In our example, we let the number of possible points of each exam could vary on an interval as in the following TABLE (TABLE IV).

**TABLE IV: Possible points in the form of intervals**

| Exams | First midterm | Second Midterm | Quizzes | Lab | Final Exam |
|---|---|---|---|---|---|
| Possible points | 10-20 | 15-25 | 5-15 | 2.5-7.5 | 25-75 |

### 4.2 Formulation

#### 4.2.1 General

With the condition that the sum of points of all exams/assignments is 100, we can formulate this problem as a mixed linear programming model

I: the set of exams/assignments

J: the set of students

m: the number of exams/assignments

N: the number of students

$l_i$ : are the lower bound on the percent of exam/assignment i $i = 1, \dots, m$

$b_i$: are the upper bound on the percent of exam/assignment i $i = 1, \dots, m$

α: is the success threshold

$a_{ij}$: percent of points attained from exam/assignment i by student j $i = 1, \dots, m;\ j = 1, \dots, N$

$y_j$=1 if the student j succeeds and 0 if he/she fails $i = 1, \dots, m$

$x_i$: percent of points (weight from the 100 points) to be allocated to of exam/assignment i $i = 1, \dots, m$

$u_j$ is the percent of the total points attained by student j after the new distribution $j = 1, \dots, N$

The mixed integer Linear program is formulated as follows:

$$\max \sum_{j=1}^{j=N} y_j$$

St

$$\sum_{i=1}^{i=m} a_{ij}\, x_i = u_j \quad j = 1, \dots, N \qquad (1)$$

$$u_j - y_j \le \alpha \qquad j = 1, \dots, N \qquad (2)$$

$$-u_j + y_j \le 1 - \alpha \qquad j = 1, \dots, N \qquad (3)$$

$$\sum_{i=1}^{i=m} x_i = 1 \qquad (4)$$

$$l_i \le x_i \le b_i \quad i = 1, \dots, m \qquad (5)$$

$$y_j \in \{0,1\} \quad u_j \ge 0 \qquad (6)$$

The objective function tries to maximize the number of successes

The first set of constraints sums the percent obtained in each exam/assignment for each student to obtain the percent of the total points attained by student j after the new distribution.

The second and the third set of constraints impose the condition that if $u_j \le \alpha$ (the percent of points accumulated by a student is less than the threshold of success), then $y_j = 0$ ; Otherwise $y_j = 1$

The fourth set of constraints dictates that the sum of points allocated to each exam/assignment is 100%

The fifth set of constraints impose the condition that each exam/assignment has a lower number and an upper number of points.

The sixth set of constraints says that either the student is a success or failure and that the collected number of points is non-negative.

#### 4.2.2 Example

I=1…,5 is the exam or the assignment number, i=1: First midterm; i=2: Second Midterm, i=5: Final exam

J=1,..,8 is the number of the student, j=1: student 1, , j=2: student2, .., j=8: student 8

$x_i$: percentage of points (weight from the 100 points) of exam/assignment i with the new allocation. For example, the $x_1$ must be between 10% and 20% according to TABLE IV

$a_{ij}$: percent of points attained from exam/assignment i by student j (the entries are in TABLE II)

$y_j$=1 if the student j succeeds and 0 if he/she fails

$u_j$ is the percent of the total points attained by student j after the new distribution

Assuming that a student getting 0.595 or more is a success and the objective function tries to maximize the number of successes

The first set of constraints in the below program gives the percent of points accumulated by a student after modification of the possible points

The second set of constraints imposes the condition that if $u_j$ is more than 0.595 then the $y_j$=1 otherwise the $y_j$ could be 0 or 1.

The third set of constraints imposes the condition that if $u_j$ <0.595 then $y_j$=0 otherwise the $y_j$ could be 0 or 1.

Both the constraints set 2 and 3 impose the condition of success and failure of student j, $j = 1, \dots, 8$.

The constraint 4 states that the sum of all the percent of possible points is 1.
The remaining constraints state that the weight of each exam/assignment obeys the condition on minimum and maximum points assigned to each of the learning measurement given by the department council.

$$\max \sum_{j=1}^{8} y_j$$

$$\sum_{i=1}^{5} a_{ij} x_i = u_j \qquad j = 1,..,8$$

$$u_j - y_j \leq 0.595 \qquad j = 1,..,8$$

$$-u_j + y_j \leq 0.405 \qquad j = 1,..,8$$

$$\sum_{i=1}^{5} x_i = 1$$

$$0.10 \leq x_1 \leq 0.20$$

$$0.15 \leq x_2 \leq 0.25$$

$$0.05 \leq x_3 \leq 0.15$$

$$0.025 \leq x_4 \leq 0.075$$

$$0.25 \leq x_5 \leq 0.75$$

$$y_j \in \{0,1\} \quad u_j \geq 0$$

After using the software of optimization Linear Interactive Discrete Optimization (LINDO) [10] , we notice that the number of successes has increased by 1 (see TABLE V); that is, the student number 3 and this is with the new distribution of points (see TABLE VI)

**TABLE V: The number of points collected with the new distribution**

| Student | Points accumulated |
|---|---|
| 1 | 62.750 |
| 2 | 63.450 |
| 3 | 59.500 |
| 4 | 47.750 |
| 5 | 47.750 |
| 6 | 45.925 |
| 7 | 49.375 |
| 8 | 61.250 |

**TABLE VI: Optimal distribution of the points among exams**

| Exams | First midterm | Second Midterm | Quizzes | Lab | Final Exam |
|---|---|---|---|---|---|
| Possible points | 10 | 15 | 5 | 7.5 | 62.5 |

## 5. RESULTS AND DISCUSSION

### 5.1. Solving the first problem

if we look at the TABLE of comparison (TABLE VII), we can notice that for some students the scores have increased and for others they have decreased and the number of successes has increased by 1; that is student 3.

**TABLE VII: Comparisons of the two scores**

| Student | Points accumulated With the initial possible points | Points accumulated With the new possible points |
|---|---|---|
| 1 | 73 | 62.75 |
| 2 | 62.05 | 63.45 |
| 3 | 59 | 59.5 |
| 4 | 47.5 | 47.75 |
| 5 | 48.25 | 47.75 |
| 6 | 45.2 | 45.925 |
| 7 | 49.5 | 49.375 |
| 8 | 61.5 | 61.25 |

### 5.2. Solving a Large scale problem

We tested our procedure on Large size problem with real data; that is, a course of 302 students and 4 exams/assignments with initial distribution of possible points as follows: 15 points for the first exam, 20 points for the second exam, 15 points for the third exam and 50 points for the final exam. Imposing the bounds on the possible points as in TABLE VIII,

**TABLE VIII: The intervals over which the possible points are distributed**

| Exams | First exam | Second Exam | Exam/Assignment | Final Exam |
|---|---|---|---|---|
| Possible points | 10-25 | 10-30 | 5-25 | 40-70 |

we programmed the problem with Cplex 12.1 [11].
The problem has 608 variables 1519 constraints, After running the program, we found that the number of successes has increased from 207 to 220; that is, we gained 13 successes from assigning the possible points as follows: the first exam 11.7 points, the second exam 10 points, the third exam/assignment 25 points and the final exam 53.3 points.

## 6. CONCLUSION

A choice of possible points' distribution could impact the success of a student in a course. I have developed a new scheme using mixed linear programming in order to eliminate the component that can influence the success of certain students. My method provides an optimal choice of weights of exams/ assignment to provide fairer grading system. Future research can concentrate on comparison of different methods as far as accuracy is concerned and applying the method in other fields such as budget allocation.

## REFERENCES

[1]. Reeves, Douglas (2008) Leading to change/Effective grading Educational Leadership **65**(**5**) pp85-87.
[2]. Wankat, P., Oreovicz, F (2015) Teaching Engineering, Second Edition Purdue University Press chapter 11.
[3]. Tinkelman, D. Venuti, E. Schain, L (2013). Disparate Methods of Combining test and assignment scores into course grades Global Perspectives on Accounting Education Volume 10, 61-80.
[4]. Wilson, I (2008) Combining assessment scores – a variable feast Medical Teacher; 30: 428–43
[5]. McLachlan, J. Marks, S (2000) scores and grades: scaling and aggregating student's assessment outcomes Medical Education 34:788–797
[6]. Douglas, R., Mislevy, J (2010) Estimating Classification Accuracy for Complex Decision Rules Based on Multiple Scores Journal of Educational and Behavioral Statistics, Vol. 35, No. 3, pp. 280–306
[7]. Vanderbei,J. Scharf, G. Marlow, D (2014) A Regression Approach to Fairer Grading SIAM REVIEW Society for Industrial and Applied Mathematics Vol. 56, No. 2, pp. 337–352
[8]. Mcnamara, J (1973). Mathematical programming applications in education planning. Socio Economic Planning Science. Vol. 7, pp. 19-3
[9]. Johnes, J (2015) Operational Research in Education, European Journal of Operational Research Volume 243, Issue 3, 16 June 2015, Pages 683–696
[10]. Linear Interactive Discrete Optimization (LINDO) www.lindo.com.
[11]. Cplex12.1<http://www01.ibm.com/support/docview.wss?uid=swg21647410>.