ES112.02/04

Fall 2003-2004 Term Project

Project #1: "Curve Grader"

1 The Project

The goal of this project is to write a computer program that takes as input the names and (numeric) grades of all students in a class, and assigns them letter grades using the "curve" method. You can find the details and requirements below.

2 Some Statistical Preliminaries

In order to describe the method of "grading on a curve", some statistical definitions must be known. Let $x_i = \{x_1, x_2, \ldots, x_{N-1}, x_N\}$ be a set of values (such as grades in a class). The *arithmetic mean* (or just the *mean*) is denoted $\overline{x}$ and is defined as follows:

$$\overline{x} = \frac{1}{N}\sum_{k=1}^{N} x_k \tag{1}$$

The *variance* is a measure of the spread of values around the mean. The variance is denoted by σ_x^2, and it is defined as below:

$$\sigma_x^2 = \frac{1}{N}\sum_{k=1}^{N}(x_k - \overline{x})^2 \tag{2}$$

The *standard deviation* is defined to be the square root of the variance and (not surprisingly) denoted by σ_x. It is the more "appropriate" measure for the spread of values around the mean since it has the same units as x_i and $\overline{x}$. Just to be explicit, the definition is:

$$\sigma_x = \sqrt{\sigma_x^2} \tag{3}$$

3 Grading on a Curve

There is not much of a secret to grading on a curve. The assumption is that there are enough many grades ($N \geq 20$) and the grades will have a "normal" (or Gaussian) distribution around the mean. (You do not have to understand what that exactly means.)

Operationally, what is to be done is as follows. First, you take the grades x_i, and calculate the mean ($\overline{x}$) and the standard deviation (σ_x). Now, we use a linear transformation on each x_i to convert it to a scale where the mean is 2.0 and the standard deviation is 1.0. This really sounds harder than it is. Simply, if you call the transformed grades y_i, they can be calculated using the formula below:

$$y_i = 2.0 + \frac{x_i - \overline{x}}{\sigma_x} \tag{4}$$

The transformed grades (the y_is) are really the grades on the scale of 4.0. On this scale, a 2.0 corresponds to the letter grade CC, 3.0 to the letter BB, and so on. However, we must divide midway between each grade to make this all work. The table below shows how the letter grades should be assigned depending on the transformed grades y_i.

Range	Letter Grade
$3.75 \leq y_i < \infty$	AA
$3.25 \leq y_i < 3.75$	BA
$2.75 \leq y_i < 3.25$	BB
$2.25 \leq y_i < 2.75$	CB
$1.75 \leq y_i < 2.25$	CC
$1.25 \leq y_i < 1.75$	DC
$0.75 \leq y_i < 1.25$	DD
$-\infty \leq y_i < 0.75$	F

4 Subjective Modification

The pure method of "grading on a curve" is only valid when the class in question is comprised of "average" students. In other words, curve grading only evaluates the *relative* performance of the students. In order to get the grades correct in an *absolute* sense, one must also consider the level of the whole class. Unfortunately, very often there is no objective method for judging the level of the whole class. The judgement is then left to the instructor.

No matter how the class level is measured, it is factored in as an additional constant to the transformed grades. If we denote the transformed and subjectively modified grades z_i, they are given by:

$$z_i = y_i + \alpha \tag{5}$$

Here, α is a constant that depends on the level of the class. The constant α is most commonly in the $[-1.0, 1.0]$ range.

Some sample values for α and their meaning for the class level is given in the table below:

α	Class Level
+1.0	Very good class of select students.
+0.5	Class of well-performing, interested students.
+0.0	Average class
-0.5	Class of bad-performing, uninterested students.
-1.0	Class consisting of proven low-performers or repeating students.

Just for completeness, the definition of z_i can also be given in terms of the original grades:

$$z_i = 2.0 + \alpha + \frac{x_i - \overline{x}}{\sigma_x} \tag{6}$$

Once the z_i are calculated, the letter grades are assigned according to the same table given in the previous section. If you examine what is going on, it should be clear that setting $\alpha = 0.0$ means assigning CC to the class average, setting $\alpha = 0.5$ means assigning CB to the class average, setting $\alpha = -0.5$ means assigning DC to the class average, and so on.

5 Project Requirements

In this project, you are supposed to write a C program that takes the names and raw grades of students plus the modification constant as input, and generates a list containing the names and letter grades of students in sorted order, and also the mean and standard deviation of the class.

The input should be a text file, containing the modification constant, and the list of students and their raw grades in the following format:

```
0.5

John Doe: 45
Jack Smith: 76
Lisa Kwok: 47
William Schmitt: 88
Vladimir Zelevinsky: 51
June Matthews: 77
George Winston: 33
Harold Peterson: 62
Dave Pritchard: 91
P.A.M. Gram: 69
```

The first line in the file will contain a single number, which is going to be the overall grade modification constant. After that, each line that is not blank, will contain the name of a student, any number of spaces, a colon (':') followed by the raw grade of the student.

Your program is supposed to read the file, ignoring any blank lines. It should do the necessary calculations, and output the list with the letter grades to another text file.

When the program is first run, it should ask the user for the name of the input file. When it is entered, it should try to read that file in. If there is an error (file can not be found, file contents can not be parsed properly) it should report the error, and ask for the input filename again. Otherwise, it should ask for the output file name. Once that is entered, it should write its output to the given filename, unless there are any errors. Then, the program should exit.

You can assume that the name of no student is longer than 79 characters and no line in the file exceeds 255 characters. You can also assume that `float` precision is sufficient for all calculations.

5.1 Hints

Since you do not know how many students are in the class, you will need to first go through the file once, counting the lines that are non-blank, except for the first line which is supposed to contain the adjustment to the transformed grades. After that, you should read the file again, this time storing the names and raw grades in appropriate structures.

5.2 Sample Output

The output of the program (which should go into an output text file) should look roughly as follows (given the input file above):

```
 1             John Doe: DC
 2           Jack Smith: BB
 3            Lisa Kwok: DC
 4      William Schmitt: AA
 5 Vladimir Zelevinsky: CC
 6        June Matthews: BB
 7       George Winston: DD
 8      Harold Peterson: CB
 9       Dave Pritchard: AA
10          P.A.M. Gram: BB
10 students total.
Mean: 63.90
Standard deviation: 18.51
Adjustment used: 0.50
```