61A Exam Data Analysis: Measure of Academic Dishonesty

Kavi Gupta

January 5, 2017

1 General Strategy

The general strategy is to look at similarities between scores of students sitting next to each other against the similarities between scores of students sitting not next to each other.

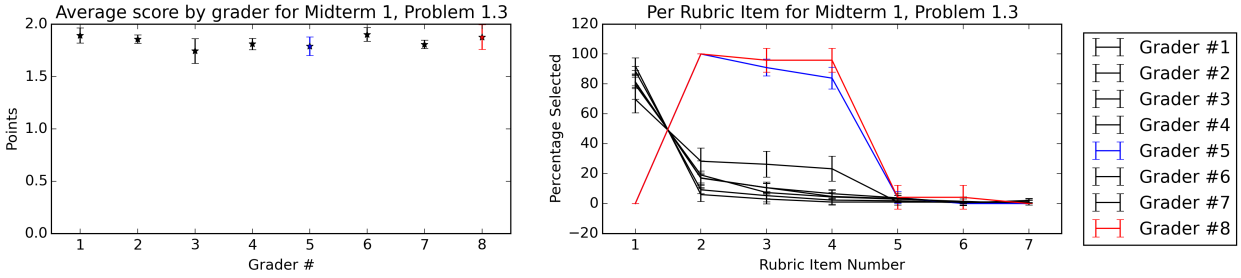
The result can be then used to estimate the number of students committing academic dishonesty using a model.

2 Confounding Factors

There are many potential confounding factors, which we control for in different ways.

2.1 Grader Differences

Some graders grade differently from each other. Taking Midterm 1 problem 1.3, we can see that graders 5 and 8 had fairly typical behavior in terms of average scores given. However, we can also see that they have a very different profile of rubric items. In this case, what happened was that these two graders were giving rubric items 2, 3, and 4, which corresponded to each of the three parts of the problem being correct, rather than giving rubric item 1, which corresponded to the entire problem being correct.



Since grading ranges tend to be at least somewhat associated with location, this might lead to students near each other having more similar profiles, artificially. We control for graders by subtracting out from each student's individual problem scores and rubric items the mean given by a grader.

Any student who had at least one problem that was either graded by a grader who had graded fewer than ten problems (not enough of a track record) or who had an abnormal pattern of rubric item grades[1].

[1] Unusualness of a grader is defined as

$$u = \sum_k \left| \frac{x_k - \mu_k}{\sigma_k} \right|$$

where x_k is the mean for the given grader of the kth rubric item, and μ_k, σ_k are the mean and standard deviation for that rubric item in general