Numerical grades as engines of inequality

Jon Zelner   Sarah Zelner

Introduction

In a recent post on the blog *Grading for Growth*, Talbert (1) argues against the use of numerical grades resulting from averaging assignment grades as a kind of category error: Rather than being numerical measurements like ones for mass, height, etc., grades are actually categories with a numerical label. These categories are ordinal, i.e. their ordering has some inherent meaning, but taking the sum or mean of the *labels* used doesn't necessarily convey useful information. In other words, most of the summary statistics we use to aggregate assignment scores have no inherent meaning in the same way that the average height of a group of people or some other summary of numerical measurements might have.

The other thing that these types of summaries do is that they typically ignore improvement over time, which of course suggests growth and effort on the part of students. As a result, a student who gets a very low grade on the first assigment of the term but a nearly perfect one on the final paper may get the same final grade as a student who is posting mediocre grades all term but exhibits essentially no improvement.

These systems necessarily advantage students who come in with a high baseline of knowledge and comfort in the classroom while disadvantaging those who may come in with a lower level of preparedness but exhibit marked improvement over the course of the term.

In this paper, we examine the imapct of different classical and progressive grading schemes on inequity in end of term outcomes. We begin with the baseline assumption that end-of-term grades should reflect growth over the course of the term rather than baseline preparedness and comfort. To assess the impact of grading schemes on inequity, we will examine the relationship between the distribution of end-of-term grades as compared to the distribution of a) beginning of term preparedness and b) individual rates of improvement over the course of the term. Grading schemes that promote equity will result in an essentially linear relationship between individual improvement and end-of-term grade, whereas those that promote inequity will largely reflect variation owing to beginning-of-term attributes rather than growth.

Methods

We will utilize a model of student knowledge and assessment performance that is analogous to a logistic regression model. We will begin with the simplifying assumption that assignment grades are perfect measures of performance, i.e. they are error-free measurements. Later on, we will complicate the notion of grades-as-measurement by including a relationship between beginning-of-term preparedness and measurement error. But beginning with this simple scenario will allow us to explore the implications of different schemes for grade inequity under even those scenarios in which assignment grades are decided upon without social bias.

Simulation Model

Each student has an individual beginning-of-term expected performance, denoted as $100 * (1 - e^{-\alpha_i})$ and a semester log-improvement ratio denoted as β_i. Values are multiplied by 100 to allow outcomes to be measured on the classical 0-100 grading scale. We assume that time is measured relative to the beginning of the term, so that $t = 0$ represents the first day and $t = 1$ represents the last day.
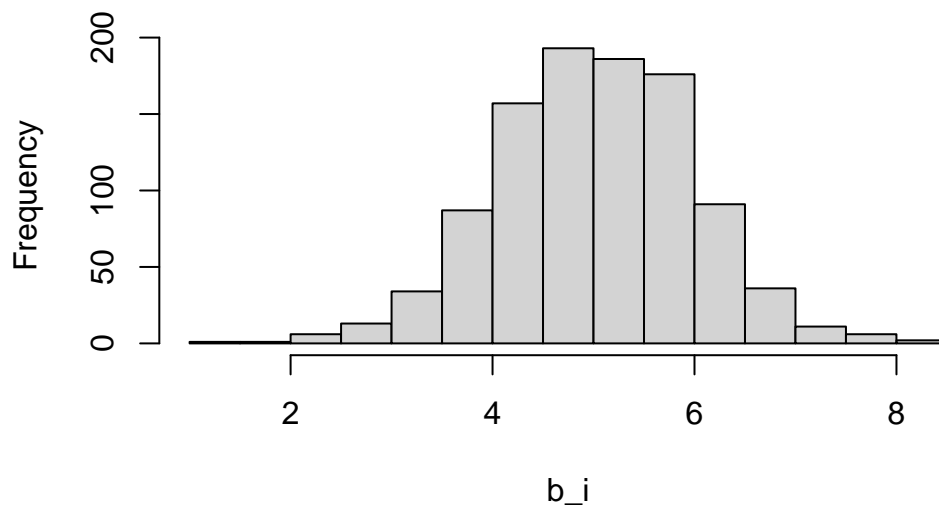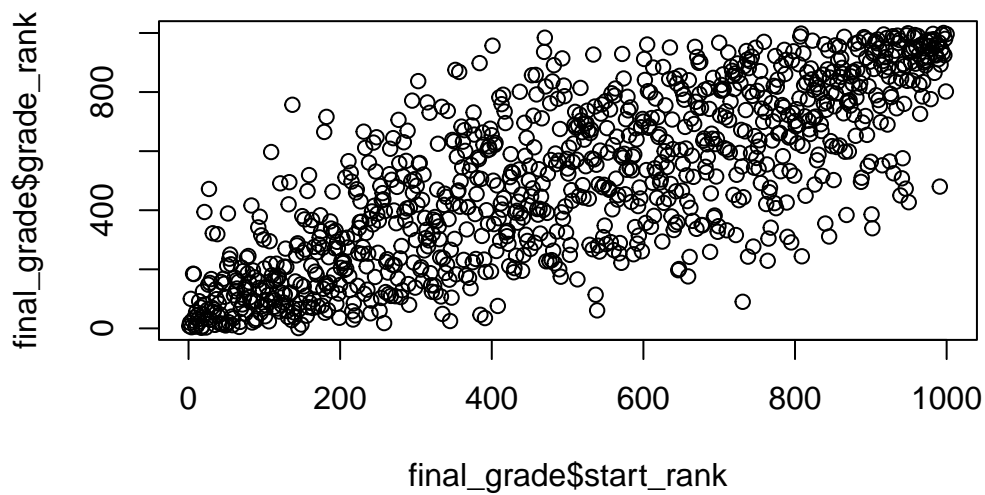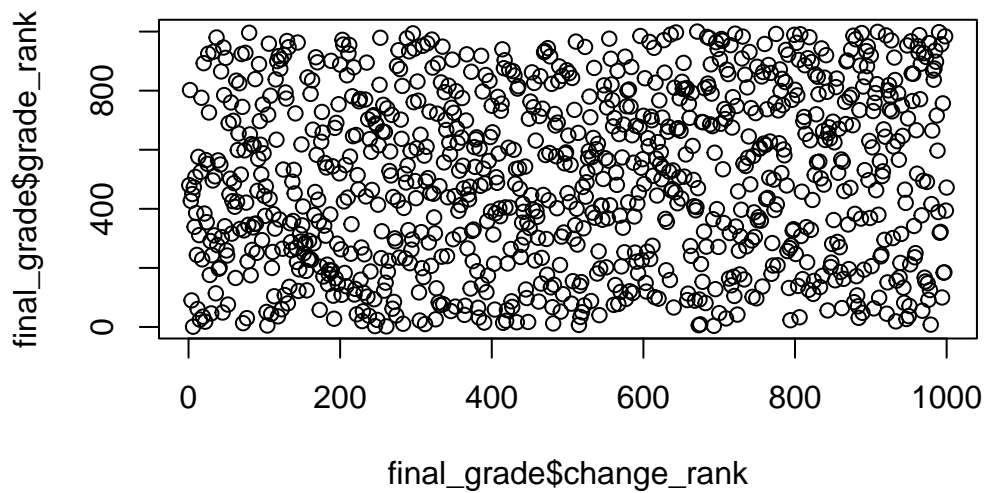
Grading Scenarios
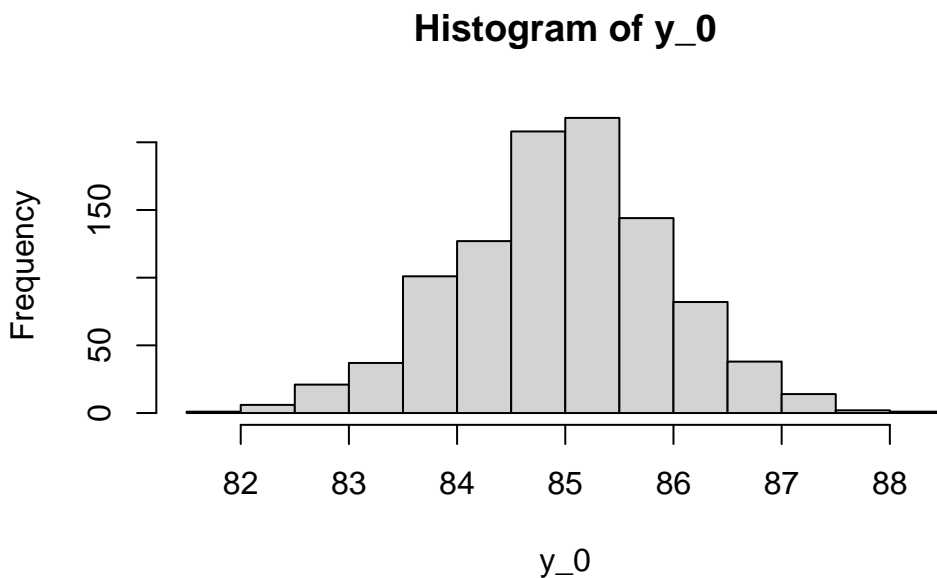
Histogram of y_0
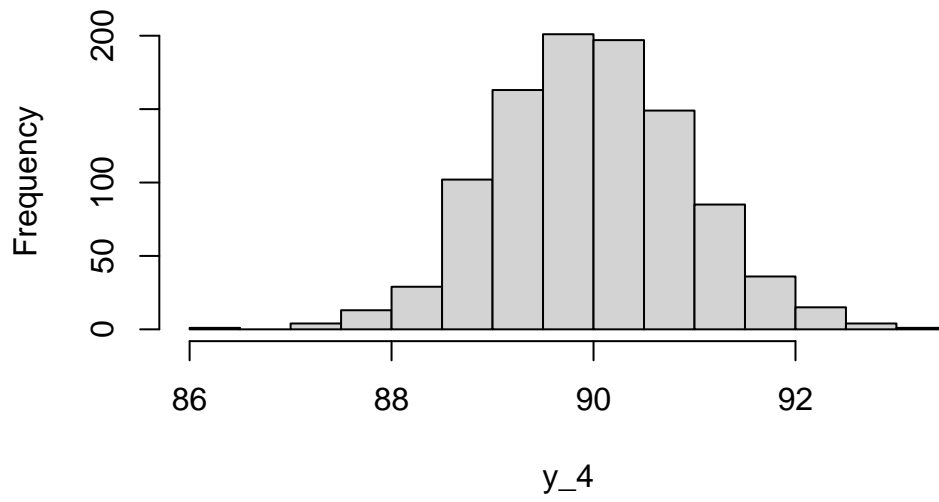
Histogram of y_4

Histogram of b_i

```
Call:
lm(formula = grade_rank ~ change_rank, data = final_grade)

Residuals:
    Min      1Q  Median      3Q     Max
-548.64 -245.48   -1.59  243.99  544.84

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 441.77362   18.16307  24.323   <2e-16 ***
change_rank   0.11734    0.03144   3.733    2e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 287 on 998 degrees of freedom
Multiple R-squared:  0.01377,   Adjusted R-squared:  0.01278
F-statistic: 13.93 on 1 and 998 DF,  p-value: 0.0002003
```

References

1.   Talbert R. When is a number not a number? *Grading for Growth.* 2022;(https://grad ingforgrowth.com/p/when-is-a-number-not-a-number). (Accessed April 26, 2022)