Estimating the Student Population GPA Increase: Consequences of Implementing a Lenient Pass/No Credit Grade Policy

Christopher Newmark[1*]

[1*]Mathematics, California Lutheran University, 60 W Olsen, Thousand Oaks, 91360, California, United States.

Corresponding author(s). E-mail(s): cnewmark@callutheran.edu;

Abstract

Pressure from the COVID-19 pandemic resulted in lenient grading policies for some Universities during the 2020 - 2021 academic year [**?**]. Certain grants stipulate reporting data on student grade-point average (GPA), and thus it is important to estimate and remove the artificial GPA increase resulting from changes in grading policies in order to understand the true effect of grant work. Moreover, administration might want to estimate the effect their grading policies will have on their student population. In this paper, we show a theoretical upper bound calculation for individual student GPA increases, as well as a simulation of a population level GPA increase. Our results show an upper bound of 33% increase with an average population GPA increase of 20% under the naive assumption of all grades equally likely. Finally, we apply the three year historical grade distribution of our university to show an average population GPA increase of 4%.

Keywords: Grade Point Average, Grade Policy, Monte Carlo Simulation, Pass / No Credit

1 Introduction

As a result of the COVID-19 pandemic, many institutions adopted compassionate grading policies. These policies often centered around the pass / no pass option, including a new option of low-pass. Additionally, students might be given more time to decide to withdraw or change their grading option going as far as allowing for retroactive changes. Assuming that students would choose one of these options if they believed they would receive a grade lower than a C, this has caused GPAs to inflate. However, many university grants report GPAs and GPA percent increases as part of their continued monitoring and funding. Thus, there is a need to removed artificial GPA inflation in order to estimate the true impact of grant work. Since we don't know what students' grades would have been had they not chosen the pass / no pass option, we don't have a control population to compare their grades. Thus, Monte Carlo simulation is a good choice to create an artificial control population. We ran two simulations - one simulation was under the naive assumption that each grade was equally likely. In the second simulation, we used university historical grade data to construct the grade distribution from which we sampled. We also assumed that students were taking between three and five classes. Finally, we assumed that every student would change one bad grade (but not necessarily their lowest) to the Pass / No Credit option if they could.

2 Theoretical GPA Estimate

We can calculate the largest GPA increase a student would receive by dropping one course as follows.
Assuming the standard mean calculation, if a student takes $'n'$ classes with grade points $x_1, ..., x_n$ and drops one class, their original GPA calculation is

$$\frac{x_1 + ... + x_n}{n} \tag{1}$$

Without loss of generality, assume they drop the course corresponding to grade x_n. Their new GPA calculation is

$$\frac{x_1 + ... + x_{n-1}}{n-1} \tag{2}$$

We now want to calculate the percent increase from their original GPA to the new GPA. We want to solve for $'y'$ in the following

$$\frac{x_1 + ... + x_n}{n} \cdot y = \frac{x_1 + ... + x_{n-1}}{n-1} \tag{3}$$

$$y = \frac{n}{n-1} \cdot \frac{x_1 + ... + x_{n-1}}{x_1 + ... + x_n} \tag{4}$$

Let $t = x_1 + ... + x_{n-1}$. Then the above equation becomes

$$y = \frac{n}{n-1} \cdot \frac{t}{t + x_n} \tag{5}$$

This equation is maximized when $x_n = 0$. In other words, the GPA increases most when a student drops what would have been an F grade. Thus, $y = \frac{n}{n-1}$.

For example, if a student taking five classes drops one of them, their maximum GPA increase is $y = \frac{5}{4} = 1.25$ corresponding to a %25 increase.

3 Results

Under the naive assumption that each grade is equally likely, our results showed that about %20 of the new GPA value was due to dropping one bad grade. To make this clear, this is not an estimate in the percent increase in GPA, but rather, an estimate of the percentage of GPA points that are due to the change in policy. Under the historical grade distribution assumption, our results showed that about %4 of the new GPA value was due to dropping one bad grade. This dramatic difference stems from the historical grade distribution being skewed toward A and B grades.

4 Python Code

Note: The code you see has been edited for formatting purposes and readability.

```
#!/usr/bin/env python
# coding: utf-8

# In[1]:


n = 100 #students
s = 100 #semesters


# In[2]:


import numpy as np
import pandas as pd
import matplotlib as mpl


# In[3]:


import matplotlib.pyplot as plt
from pandas.plotting import table


# In[4]:
```

```
#importing historical grade counts
hist_grade_dist = pd.read_csv('TUG_Grade_Distribution.csv')


# In[7]:

letterGrade = hist_grade_dist.iloc[0:12,:]


# In[9]:


gradePoint_dict = {'A':4,'A-':3.7,
                    'B+':3.3,'B':3,'B-':2.7,
                    'C+':2.3,'C':2,'C-':1.7,
                    'D+':1.3,'D':1,'D-':0.7,
                   'F':0}


# In[10]:
x = letterGrade['Grade'].map(gradePoint_dict)
letterGrade['gradePoint'] = x

# In[12]:

countTotal = letterGrade['Count'].sum()
x = letterGrade['Count'].map(lambda s: s / countTotal)
letterGrade['CountPercent'] = x


# In[14]:

x = letterGrade.loc[:,'CountPercent'].cumsum()
letterGrade['CountPercent_Cumu'] = x


# In[21]:


sim_list = [] #hold arrays of students
x = pd.DataFrame(columns=['Average GPA No Drop',
                          'Average GPA with Drop'],
                 index=['Semester'])
mean_frame = x
#keeps track of summary statistics
```

```
#through the semesters

# In[23]:


def grade_choice():
    return np.random.choice([0,1,2,3,4])


# In[24]:


points = list(letterGrade['gradePoint'])
percent = list(letterGrade['CountPercent'])
#points , percent


# In[25]:


def historical_grade_choice(points,perc):
    return np.random.choice(points,p=perc)


# In[26]:


def num_classes():
    return np.random.choice([3,4,5])


# In[27]:


def drop_badGrade(ls):
    if any(ls < 2):
        y = [x for item in
             np.argwhere(ls < 2).tolist()
             for x in item]
        drop_index = np.random.choice(y)
        return np.delete(ls,drop_index)
    else: return ls
```

```
# In[28]:


arr = []
test = []
sim_list.clear()

#number of semesters / simulations
for i in np.arange(s):
    for _ in np.arange(n): #number of students
        for _ in np.arange(num_classes()):
            #arr.append(grade_choice())
            arr.append(
                historical_grade_choice(points,percent)
                )

        sim_list.append(np.array(arr.copy()))
        arr.clear()

    test.append( [
     np.mean([np.mean(student) for
              student in sim_list]),
     np.mean([np.mean(drop_badGrade(student)) for
              student in sim_list])
    ])

    sim_list.clear() #clear sim_list every semester

#index=pd.RangeIndex(start=1, stop=s, name='Semesters')
x = pd.DataFrame(test,
                 columns=['Average GPA No Drop',
                          'Average GPA with Drop'])
mean_frame = x

# In[30]:


mean_frame.index_name = 'Semesters'


# In[32]:

x = mean_frame.loc[:, 'Average GPA with Drop']
y = mean_frame.loc[:, 'Average GPA No Drop']
mean_frame['Differential'] = x - y
```

```
# In[33]:

#How much of percentage is due to artificial increase?
mean_frame['Percent Increase'] = (x - y) / y * 100


# In[35]:


mean_frame.loc[:, 'Percent Increase'].mean()
```

5 Methods

We wanted to compare the GPA of student population to the GPA of those same students if they had dropped a bad grade. We randomly assigned between three and five classes to each of one-hundred students. Then, in two cases, we assigned those classes grades.

In the first case, we assigned grades F,D,C,B, and A with equal likelihood. The average GPA for this cases was 2.0. Then we randomly dropped a D or F grade and recalculated the average GPA. This was one semester of simulation. We then repeated the above process for one-hundred semesters. Finally, we took the difference between the average GPA with and without dropping the grades and looked at the ratio of this increase to the increased average to determine the percent of average GPA increase for each semester that was artificial. Then we averaged those values.

In the second case, we followed the same procedure assigning instead grades of F,D-,D,D+,C-,C,C+,B-,B,B+,A-, and A with likelihood equal to the historical grade distribution.

6 Conclusion

Originally, our grant was looking at GPA percent increases over three years. We saw good increases that could be attributed to our grant work, but there was an asterisk next to the third year which saw an increase that was marginally higher than the previous two years. We sought to answer the question, how much of this percent increase is due to our work versus the artificial increase of the grade policy? The simulation showed that only a small percentage of percent increase was due to the grade policy which meant that a majority of the percent increase we saw could be attributed to our grant work which is vital information for reporting purposes.
Other refinements can be made to this simulation. For example, we might

assume that not all students will drop a bad grade, because they wait too long to do so. Another refinement example is that there may be a good amount of students who are part-time or take more than five classes.
Finally, there are broad implications for universities which are planning or have already implemented compassionate grading policies. The effect the university may see will depend on their historical grade distribution.

Acknowledgments. Thank you Dr. Delil Martinez for your guidance on this paper. You are above the mean!