STATISTICAL METHOD TO PREDICT GRADUATION WITH DISTINCTION AT COLORADO STATE UNIVERSITY

Abstract ID: 1806

William S. Duff, PhD, PE
Colorado State University
Fort Collins, Colorado

Jannette Pérez-Barbosa, MS, PE
Universidad del Turabo
Gurabo, Puerto Rico
Graduate Student, Colorado State University

Abstract

A statistically based methodology was developed at Colorado State University to create a fair Graduation with Distinction honor designation process. This methodology considers the variation of class rankings over time using time indexed control charts. These statistical process control tools track GPA cutoff values for Summa, Magna and Cum Laude honors designations and signal the need for changes in the cutoff values. Analysis was performed for the last ten academic years considering both spring and fall graduation semesters with graduation candidates in all of Colorado State University’s undergraduate colleges. This methodology is currently being studied to further develop the statistical process control methodologies in the context of extreme value statistics. This research reports on the experience with the methodology and anticipated improvements in the methodology.

Keywords
Engineering Education, Control Charts, Cumulative Sum, Extreme Value Statistics

1. Introduction

Universities around the world award honors to students completing their studies as recognition of their academic excellence. These recognitions are generally reserved for a small percent of the graduating class and can be mentioned by honorees in their resumes or other professional documents. Colorado State University (CSU) awards three tiers of Academic Distinction: Magna, Summa and Cum Laude in all of its nine Colleges. The percentile cutoff values for the awards are the top 1% (Summa Cum Laude), the next 3% (Magna Cum Laude) and finally the next 6% (Cum Laude) [1]. The required grade point averages (GPAs) and total number of awards in each category will depend on the size of the college.

The process of identifying potential students to be awarded these honors (candidacy) is usually made before the graduation date (at the end of each semester). Based on the methods used before 2004 the notification did not guarantee actual distinction at commencement. Dr. William Duff has been working on establishing a methodology that facilitates the awarding of these honors each year in a way that increases the fairness of the process. Initially having just one GPA for each of the percentiles university wide was considered. However, it became apparent that there are significant grading differences between CSU’s colleges that required GPA breakpoint values for individual colleges.
A statistical procedure to determine the desired GPA breakpoints was created in 2004 and is currently in use at the university. This approach includes the statistical concepts of: central tendency and dispersion measures, time phased control charts (CUSUM) and proportion control charts (P-charts). These allow monitoring to detect if there are statistically significant changes in the GPA distributions to qualify for an honor as compared to historical data. This procedure is being evaluated to review its performance after ten years.

Data from the last ten years will be utilized to compute the target GPA for each honor level. Statistical comparisons will be made in order to determine if there is significant variation from year to year in the actual GPAs as compared to the predetermined cutoff GPA’s. An example will detail the procedure to obtain the honor cutoff values and compare them to actual GPAs to assess the significance of changes over the investigation period.

The paper will also review some of approaches of other universities regarding their graduation with distinction procedures in order to compare with CSU’s current procedure. Other research described will include weighted control charts and extreme value statistics methodology. The last part of the paper will discuss the next steps related to the future development of refinements to the current calculation method based on the use of extreme value statistics.

2. Problem Definition

The policy for Graduation for Distinction Candidacy allows outstanding students that comply with the requirements to be awarded a distinction based on a fair methodology using established by time indexed control chart (CUSUM) GPA breakpoints. This was established by Dr. William Duff in 2004 with initial calculations and statistical analysis to support the method. As part of the implementation of the methodology, the minimum GPA’s suggested must be reviewed every four years to determine if the percentage of students graduating with distinction has statistically deviated from the target percentages of 1%, 4% and 10% of graduates. The four year review cycle provides entering freshmen with a fixed target for achieving a distinction level. The initial analysis also supports a four year review cycle as there is normally not significant variation of the GPA during that review period. After the new Graduation with Distinction Method was established, the university has performed the re-calculation of the GPA’s every four years. This method is described along with the current minimum GPA’s for distinction in the University’s General Catalog [1].

After a decade of using this methodology there is a renewed effort to develop a more refined statistical process control methodology that can more accurately provide GPA breakpoints required for a fair determination of honors breakpoints by applying the concepts of extreme value statistics. The event of awarding honors to the upper tier of students that represent a small part of the student population can be seen as the analysis of attributes within the context of extreme value statistics.

3. Literature Review

American universities award honors for academic distinction. These honors are usually referred as Summa, Magna and Cum Laude and a student needs to obtain a certain GPA for receiving the honor. Depending on the academic institution, the honor could be awarded to any number of candidates with a certain GPA [2] or based on the academic ranking of the student [1 and 3]. For CSU the awards consider the top 1% (Summa Cum Laude), the next 3% (Magna Cum Laude) and finally the next 5% (Cum Laude) [1]. A similar procedure was observed at the University of Idaho [3], Harvard [4] and Stanford [5]; with the latter university awarding honors to the top 15% of the graduating class. Only in the case of CSU and Idaho are the cutoff GPA’s calculated individually for each of the academic colleges.

The establishment of the current methodology required for utilization of statistical process control (SPC) tools involve two important control charts, one describing attributes and one describing variables. The use of the attribute p-charts [6] in this case is derived from the original application of fraction non-conforming control charts. In this case, the students that have achieved an academic honor will be viewed as non-conforming from the norm. This fraction will change from year to year due to different numbers of students graduating each year and different numbers of students awarded an honor every year. If the percentage of honors is not subject to assignable causes, such as grade inflation, from year to year, the control chart will remain in statistical control.
In the case of the time weighted control charts, various types are provided in the literature. For the calculations performed in this research, Cumulative Sum or CUSUM control charts were selected. According to Montgomery [6], these charts are very sensitive to small changes and will be able to detect small deviations from a target value.

An additional aspect to consider in this study is the application of the concept of extreme value data to the distinction assignment process and the estimation of parameters for that data. Extreme values can be described as data that lies in a tail of a distribution [7] such as lowest temperatures, annual minimal discharges as well as events with large extreme values [8]. Hernandez and Johnson [8] present a transformation to the exponential distribution for extreme data. Depending on the distribution to be utilized the parameters such as variance and covariance can be mathematically estimated. Reynolds and Stoumbos [9] present various cases of application of the CUSUM control chart for non-normal and extreme value data. The variation of the data is also a concern that some authors have considered, namely using the concept of Self Starting Control Charts [10]. There are additional methods to establish control charts for extreme values, including derivations from the moment estimation formulas [11]. Based on the CSU data, the awarded honors are part of the upper tail of a distribution and are considered extreme value data relative to the overall GPA’s obtained by all graduates.

4. Methodology

The initial statistical analysis performed in 2004 started with a ten year feasibility study and worked with 24 cases: the three academic honors times CSU’s eight colleges. The new methodology established breakpoints for each college to determine distinctions at graduation based on the SPC analysis performed at that time. Once implemented in 2004, the procedure specified a periodic review every four years to determine if there were significant changes in the target GPAs that require changing these values in the Academic Catalog. The method was deployed and staff at the registrar’s office were assigned responsibilities in performing the analysis for each graduation period.

A detailed analysis of the initial methodology is shown in these steps:

a. Calculate the percentages of graduates awarded an honor for each of the 24 cases.

b. Determine the standard deviation for each case.

c. Use retrospective data (1994 to 2003) to determine annual percentages of distinctions per each case.

d. Pooled data per colleges based on number of students

e. Compute targets and standard deviations for step d

f. Setup Cumulative Sum (CUSUM) control charts to detect a shift of 1.5%, 1.0% and 0.5% or more shifts from the target values for each of the cases.

g. Use proportion control charts (p-charts) to monitor the percentage of awards by case in the study period

h. If the process is in control, any changes from period to period are not statistically significant and the recommend target GPA’s can be considered applicable over the subsequent four years for a fair graduation with distinction process.

Calculations and statistical analysis considering the method will require periodic review and values will only need to be changed if there is a statistically significant variation per each review period. A relatively small shift in the difference of the mean GPA based on a target value can be detected by the CUSUM control charts. The process was seen to be sensitive to shifts of 0.5% for each GPA.

A current review using the last seven years of data (2007 to 2013), presented next, replicates the 2004 approach. The current review will replicate the steps shown previously and determine if special causes are present. The new target GPAs for each of these cases, if any of the process not in statistical control due to these special causes, will be recommended. The information utilized will also serve as a base for further development of the research topic.

5. Computing the Graduation with Distinction Target GPA’s

The calculations utilized to determine the minimum GPAs to receive academic distinction used for each of the Colleges at CSU are shown next. The initial analysis performed utilized data from 2007 thru 2013 and consolidates Fall, and Spring terms. For these calculations, the target GPA’s utilized by the university were based on the academic catalog with a review of the target GPAs in the year 2010. These values are still the breakpoints required for graduation with distinction at the university. For the already awarded distinctions, calculations were done to
compute the percentage of students that received each honor. It can be seen that only about 10% of students indeed received these recognitions as related by the policies in the University catalog.

Table 1: Summary of all CSU’s academic awards (2008-2012)

<table>
<thead>
<tr>
<th>All graduates 2008-2012</th>
<th>Awarded Honors</th>
<th>% of Class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cum Laude</td>
<td>Magna Cum Laude</td>
</tr>
<tr>
<td></td>
<td>1417</td>
<td>676</td>
</tr>
</tbody>
</table>

When considering the study, the number of honors awarded to the total class is also very close to the catalog 10% target. Also, as can be seen in table 1, the catalog requirement of top 1%, 4% and 10% of class is represented adequately. Table 2 shows the awarded distinctions for the year 2012 (data for 2012 included all academic semesters at the time of publishing this article) including all academic semesters.

Table 2: Academic distinctions for CSU, 2012

<table>
<thead>
<tr>
<th>College</th>
<th>Distinction</th>
<th>Total Graduates</th>
<th>% Of class awarded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cum Laude</td>
<td>Magna Cum Laude</td>
</tr>
<tr>
<td>Agricultural Sciences</td>
<td></td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Applied Human Sciences</td>
<td></td>
<td>79</td>
<td>47</td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>Engineering</td>
<td></td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>Health and Human Sciences</td>
<td></td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Liberal Arts</td>
<td></td>
<td>104</td>
<td>47</td>
</tr>
<tr>
<td>Natural Sciences</td>
<td></td>
<td>39</td>
<td>17</td>
</tr>
<tr>
<td>Veterinary Med and Biomed</td>
<td></td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>College of Natural Resources</td>
<td></td>
<td>12</td>
<td>11</td>
</tr>
</tbody>
</table>

For these awards, it was determined if the percentage of honors in each category was changing over time due to potential grade inflation. Since the data is discrete and describes an occurrence/no occurrence; this is characterized by the Binomial Distribution. The average proportion of the class that receives a particular award will be the center line of the control chart. If there a significant difference between the proportion of honors in a graduating class, the upper and lower control limits would be far from the center limit showing a larger dispersion. As seen on figures 1-3 there does not appear to be that particular tendency in the percentage of students that receive an award in this particular year. Once all awards are counted, the percentage of recipients in each category will be determined and evaluated for statistical control using p-charts.
The p-chart is able to show the proportion of students awarded an honor regardless of class sizes. The variation in the upper and lower control limits indicate different sizes of graduating classes. For the three academic distinctions, the proportion of the class earning any of these distinction levels remains within the control limits with no trends or patterns. The proportion of students receiving a graduation with distinction honor has not increased or decreased. Being in statistical control is confirms that the percentage of graduating class recipients is statistically the same as the percentages of students to be recognized according to the Academic Catalog. Statistically, the distinctions awarding process is consistently fairly awarding distinctions only to the legitimately worthy students.
The next graphs will show the average GPA for each type of honor and will be able to determine if a significant shift is present that moves the process out of statistical control. From 2010 thru 2013, a new set of target values was implemented for the cutoff GPA’s. Analysis for all these periods is shown next with the relevant control charts for the review period 2010-2013. When considering the average GPA in year 2010 and beyond for the Cum Laude level, it is shown to be out of statistical control, with a small shift in the average GPA apparent. Based on the methodology, this would require the recalculation of the average GPA for the academic catalog using the last four academic periods.

With the 2010 and later data it appears that the cutoff value was set too low, with all averages being over the control limit. This suggests that a calculation to adjust then average GPA back up could have been made without waiting four years. The same conclusion can be seen for the Magna Cum Laude level, the average GPA shift above the target value in an interval sooner than four years (see figure 5).
The only control chart that does not show any additional changes being required is the Summa Cum Laude seen in figure 6, as it stays within the control limits for the entire study period. This can be attributed to the very small target sample size for this group, with many students achieving the 4.00 target.

**6. Next Steps**

With these findings, and with all the data from 2013, it is being recommended that the cutoff values for all awards in all CSU’s colleges be reviewed and changed. Subsequent calculations for these values will consider the current method with the authors looking for a new and automated application that takes into consideration the extreme value distribution of this data. The new target GPAs to be used for the next four academic years in each of these cases will be also calculated. These will be reviewed with CSU’s staff in the Registrar’s office in the following months.

A more in depth Literature Review will be performed to further develop additional background knowledge regarding extreme value statistics and SPC. The distinction data will be fitted to one or more extreme value distributions. The new method will also try to determine if using the mean as a measure of central tendency is appropriate.
Acknowledgements
I would like to thank Dr. William Duff for his support in performing this research and his academic advice, support and motivation that have enabled me to continue in the Systems Engineering doctoral program at CSU. Additional acknowledgment is given as well to the CSU Registrar’s office for their flexibility to meet and share information. I am also grateful to the faculty, staff, administrators and students at Universidad del Turabo in Puerto Rico for their continuous support toward the completion of my doctoral degree.

References