Analysis of Exertion with Relative Grading Mechanism in Academics Using Computational Technique

Nagaraju Dasari¹, Mahendra Parihar² and Mohiddin Shaw Shaik³

¹Department of Mathematics and Statistics, Manipal University Jaipur, Jaipur-303007, India
²Mukesh Patel School of Technology Management and Engineering, SVKM’s NMIMS University, Mumbai, India
³Department of Mathematics, Narasaraopeta Engineering College, Narasaraopet-522601, India

E-mail: dasari.nagaraju@gmail.com

Abstract. Evaluation process is an integral part of academic system in order to monitor the performance of students from many perspectives. Since, recent past lots of methodological changes has been taken places in evaluation process from traditional one. Many experts from concerned domain claims the superiority of one method over other but still dilemma continue with reference to optimum or appropriate method. In this connection even very recently being used to a maximum extent in Indian context too is the “Grade Method”. However, even “Grade Method” is also not completely free from certain controversies. That is, though grades were initially meant to serve various pedagogical purposes, grading systems remain controversial and hotly debated when those doesn’t give recompensing results. Some argued that grades are psychologically harmful. Thus, given certain limitations of existing methods in some respect, the current study has been undertaken in order to address few issues associated with minimum passing grade and cut-off mark for highest grade - resulting out of relative grading. Therefore, in this article, through mathematical computation we propose acceptable ranges for the parameters mean and standard deviation - respectively to have appropriate mechanism and hence can make evaluation process more effective and applicable in various conditions.

1. Introduction

Assessments can broadly be classified as Formative and Summative. Assessment happening during the learning is formative or in other words, assessment FOR learning is regarded as formative. Whereas assessment happening at the end of the learning is summative or assessment OF learning is summative. Both assessments are very much needed for effective teaching and learning process. The existing literature [1][2][3] endorses that there are different types of formative and summative assessments. Assessment for learning is a process through which both the instructor and the learner would be benefitted. A proper assessment automatically facilitates Instructor to slightly tweak the approach of delivery (if required) where a student could adjust his or her learning according to the delivery of the instructor. In order to get fruitful summative results at the end of the learning, every instructor should give more emphasis on formative assessment because it happens during the teaching learning process. In an attempt to bring proper formative assessment in teaching learning process, Prashanti and Ramnarayan [4], have proposed ten maxims for its effective implementation.
As Jonathan D. Kibble [5] suggested few assessment methods for teachers while assessing students’ performance summatively. This includes identifying suitable assessment methods that meets already defined learning outcomes for a given course of delivery, timely reviewing adopted assessment methods and if required taking necessary steps in fine tuning the process and methods of assessment.

One should realize that grading of students’ performance is any course on relative grounds is an application of normal distribution under probability theory of Statistics. In fact, a proper assessment in any course would result in fair distribution of grades among the student group based on their relative performance. As per existing theory of normal distribution [6] about 68% of the population falling within one standard deviation and about 95% of the population falling within two standard deviations around the mean of considered population, shall be accepted as proper relative distribution of data or grades interms of assessing parameter for students.

2. Literature Review

Academic assessment along with appropriate evaluation method always gives fruitful results. Further, literature at various levels indicating that given particular time specific method, academic community kept on using despite of certain inherent limitations. Thus, in order to find out relevant information already exist related to study area undertaken and identify gaps, the current study has critically reviewed limited but relevant literature which can provide some basis for the study undertaken and also helps in developing appropriate methodology for the conduct of the study.

Malini and Heidi [7] study was on identifying proper rubrics in higher education emphasizing on quality and accountability associated with grading. Azmat and Iriberri [8] studied the impact of providing reward points based on individual performance in uplifting relative performance of the population or employees of any organization. Erin et.al [9] have proposed survey based rubrics in evaluating scientific papers. Andreoni and Brownback [10] through their statistical significant analysis stated that auction at the larger scale brings considerably greater rank-correlation between ability and effort in the population. Abadie et. al [11] study suggests that while clustering process of sampling and mechanism in assignment, researchers should not adjust standard errors during clustering process. Brownback [12] study reveals that students of lesser ability and slow in learning could turn in to better performers when they are in a larger group of cohorts.

Campos-Mercade and Wengstrom [13] had studied the short and long-run effects of performance threshold incentives on academic performance. In line with standard economic theory, it was found that threshold incentives improve the performance of students whose grade point average (GPA) would have been just below the threshold GPA. Further their work suggests that threshold incentives can have unexpected negative consequences on students’ academic performance. If students are overconfident about their ability, threshold incentives may damage their self-confidence and thereby harm their future performance. Shinya et. al [14] study exhibits that with increased number of examinations and mechanism of continuous feedback results the consistence further brings the improvement in students’ performance in the environment of relative grading.

In addition, by considering students sample data from different countries and institutions or organizations, few researchers [15][16][17][18][19][20][21] have analytically studied the impact of Grade Point Average (GPA) of Science, Technology, Engineering and Mathematics (STEM) courses over the other courses of study, relative grading has got no significance with respect to gender of the person, role of incentive mechanism on relative performance among a group of students, employee’s performance, children behavior.
3. Objectives of the Study

3.1. To analyse an existing academic assessment system and issues associate to it with reference to evaluation process.
3.2. To develop a mechanism in order to make evaluation process more effective and applicable in various conditions.
3.3. To examine acceptable ranges for the parameters of Gaussian distribution, mean and standard deviation respectively.
3.4. To recommend guidelines on the basis of the study for appropriate assessment and evaluation.

4. Methodology

For the study undertaken a structured approach has been adopted in order to meet the requirement of the study. The methodology adopted include literature review connected to current study. By using primary and secondary data gathered:
4.1 We first analyse the data with different mean values and standard deviations further observe the cut-off marks with reference to adopted relative grading formula.
4.2 With the given mean, we identify standard deviation intervals which suits to our anticipated cut-offs.
4.3 With graphical representation for anticipated range to cut-offs, we identify most ideal region in which mean and standard deviation should be lying in.

5. Results and Discussion

Though, it is expected auto fit of the grades under the normal curve, it should be noted/agreed that no grading policy can give effective and efficient results until unless the proper monitoring of the questions papers and evaluation scheme been followed across all the sections happens at prior.

Even with the relative grading, depending on the standard of the question paper and liberal evaluation may lead in facing difficult situation if the scoring goes higher side. Practically it is very difficult for any management to assess the delivery of the teacher and knowledge gained by the students on daily basis for each course in which he/she taken to study, of course it is not logical as well. It is largely on the basis of assessing the question paper, evaluation pattern that the teacher has followed.

Let us imagine a situation where in a course to be taught by a group of faculty members and at the end, all the students are to be graded relatively. We know each faculty has a choice to choose his/her own way of teaching a course and they may adopt different methods of evaluating the students’ performance. Therefore there would be variation in assessment which will leads to a particular group of students scoring very high marks and some group of students scoring very less marks which intern results particular group getting highest grade and whereas others getting subsequent lower grades. Hence it is in this context, it is very much essential for the management to see that grading has been done holistically and it is justifiable to the larger community of students’ to the extent possible.

As it was mentioned in the introduction, the relative grading is relatively new to the most of the institutions across the globe for them to implement and most importantly it is very much new to students and their parent community to realize the advantage of it over the traditional way of assessing the performance. Especially, it will have adverse implications if a student doesn’t get a highest grade even after scoring 90-95% marks in a given course or even after scoring 55-60% marks getting least passing grade/fails in a course through this relative grading. Which demands high level of intensity, before results get declared. Though it is possible for the management to manually check and moderate the grades which were calculated relatively irrespective of the marks secured by the students based on the
mean and standard deviation of the scores. Scatter plot and cluster analysis, curve fitting are few manual methods through which this difficulty can be addressed, but it is practically tedious task and time consuming for the administrators to monitor this activity, especially when the number of courses offered and number of students are more.

There are different learning management systems in place for most of the institutions which have adopted this relative grading system. But as of our knowledge there is no available software application for the administration to track these cases which are falling under outliers category/positively or negatively skewed grading curves. In this article we are introducing an algorithm/formula through which software application may further be developed so that there would be popup message be communicated to the administrators to intervene and take remedial steps before results get declared.

With an anticipation that passing grade range to lie between 25-35 and for highest grade range to lie between 77.5-90 and by setting cutoff mark for passing grade as mean \(-1.5\) (standard deviation) and cutoff mark for highest grade as mean \(+1.5\) (standard deviation), following analysis given in Table 1, has been done for the mean score ranging from 50 to 60 and standard deviation ranging from 15 to 20 in discrete steps.

| Table 1. Cut-Off Marks with Given Mean and Standard Deviation |
|------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| M                | a = 15  | b       | A       | a = 16  | b       | a = 17  | b       | a = 18  | b       | a = 19  | b       |
| sd = 15          | 27.5    | 22.5    | 26      | 74      | 24.995  | 75.005  | sd = 18 | 25      | 77.5    | 23      | 87.5    | 20      |
| 50               | 27.5    | 72.5    | 26      | 74      | 24.995  | 75.005  | 24.5    | 75.5    | 23      | 77      | 21.5    | 78.5    | 20      |
| sd = 16          | 31      | 76.5    | 29      | 77      | 27.995  | 78.005  | 27.5    | 78.5    | 26      | 80      | 24.5    | 81.5    | 23      |
| 51               | 28.5    | 73.5    | 27      | 75      | 25.995  | 76.005  | 25.5    | 76.5    | 24      | 78      | 22.5    | 79.5    | 21      |
| sd = 17          | 29.5    | 74.5    | 28      | 76      | 26.995  | 77.005  | 26.5    | 77.5    | 25      | 79      | 23.5    | 80.5    | 22      |
| 52               | 30.5    | 75.5    | 29      | 77      | 27.995  | 78.005  | 27.5    | 78.5    | 26      | 80      | 24.5    | 81.5    | 23      |
| sd = 18          | 31.5    | 76.5    | 30      | 78      | 28.995  | 79.005  | 28.5    | 79.5    | 27      | 81      | 25.5    | 82.5    | 24      |
| 53               | 32.5    | 77.5    | 31      | 79      | 29.995  | 80.005  | 29.5    | 80.5    | 28      | 82      | 26.5    | 83.5    | 25      |
| sd = 19          | 33.5    | 78.5    | 32      | 80      | 30.995  | 81.005  | 30.5    | 81.5    | 29      | 83      | 27.5    | 84.5    | 26      |
| 54               | 34.5    | 79.5    | 33      | 81      | 31.995  | 82.005  | 31.5    | 82.5    | 30      | 84      | 28.5    | 85.5    | 27      |
| sd = 20          | 35.5    | 80.5    | 34      | 82      | 32.995  | 83.005  | 32.5    | 83.5    | 31      | 85      | 29.5    | 86.5    | 28      |
| 55               | 36.5    | 81.5    | 35      | 83      | 33.995  | 84.005  | 33.5    | 84.5    | 32      | 86      | 30.5    | 87.5    | 29      |
| sd = 21          | 37.5    | 82.5    | 36      | 84      | 34.995  | 85.005  | 34.5    | 85.5    | 33      | 87      | 31.5    | 88.5    | 30      |

- Least passing score
- Least score to get highest grade
- Mean score
- Standard deviation

Through our analysis it tells that in a given course if the mean score falls broadly inside the interval (55, 60) and standard deviation approximately inside (16.67, 20) nothing to worry about the grading and its moderation because it produces a reasonable and fair grading with passing mark ranging from 25% to 35% and top score ranging from 77.5% to 90% to award least passing grade and highest grade in a course respectively and hence it doesn’t require any intervention from the authorities.

With reference to the following Table 2, the extrapolated range of the mean lying between (55, 60) and standard deviation lying between (16.67, 20) still be considered to be in acceptable range, but not mean going beyond 62.5 and lesser than 51.25.
Table 2. Mean v/s Interval of Standard Deviation

| Mean (m) | Standard deviation (σ) range | σ Interval for passing grade cutoff with $25 \leq m - 1.5(σ) \leq 35$ | σ Interval for highest grade cut off with $77.5 \leq m + 1.5(σ) \leq 90$ | Intersection of σ intervals |
|----------|-----------------------------|-------------------------------------------------|-------------------------------------------------|-----------------------------|
| 51       |                             | [10.67, 17.34]                                   | [17.67, 26]                                     | ϕ                           |
| 51.25    |                             | [10.84, 17.5]                                    | [17.5, 25.84]                                   | 17.5                        |
| 51.5     |                             | [11, 17.67]                                      | [17.33, 25.67]                                  | [17.34, 17.67]              |
| 52       |                             | [11.34, 18]                                      | [17, 25.34]                                     | [17, 18]                    |
| 52.5     |                             | [11.67, 18.34]                                   | [16.67, 25]                                     | [16.67, 18.34]             |
| 53       |                             | [12, 18.67]                                      | [16.33, 24.67]                                  | [16.34, 18.67]             |
| 53.5     |                             | [12.34, 19]                                      | [16, 24.34]                                     | [16, 19]                    |
| 54       |                             | [12.67, 19.34]                                   | [15.67, 24]                                     | [15.67, 19.34]             |
| 54.5     |                             | [13, 19.67]                                      | [15.34, 23.67]                                  | [15.34, 19.67]             |
| 55       |                             | [13.34, 20]                                      | [15, 23.34]                                     | [15, 20]                   |
| 55.5     |                             | [13.67, 20.34]                                   | [14.67, 23]                                     | [14.67, 20.34]             |
| 56       |                             | [14, 20.67]                                      | [14.34, 22.67]                                  | [14.34, 20.67]             |
| 56.5     |                             | [14.34, 21]                                      | [14, 22.34]                                     | [14.34, 21]                |
| 57       |                             | [14.67, 21.34]                                   | [13.67, 22]                                     | [14.67, 21.34]             |
| 57.5     |                             | [15, 21.67]                                      | [13.34, 21.67]                                  | [15, 21.67]                |
| 58       |                             | [15.34, 22]                                      | [13, 21.34]                                     | [15.34, 21.34]             |
| 58.5     |                             | [15.67, 22.34]                                   | [12.67, 21]                                     | [15.67, 21]                |
| 59       |                             | [16, 22.67]                                      | [12.34, 20.67]                                  | [16, 20.67]                |
| 59.5     |                             | [16.34, 23]                                      | [12, 20.34]                                     | [16.34, 20.34]             |
| 60       |                             | [16.67, 23.34]                                   | [11.67, 20]                                     | [16.67, 20]                |
| 60.5     |                             | [17, 23.67]                                      | [11.34, 19.67]                                  | [17, 19.67]                |
| 61       |                             | [17.34, 24]                                      | [11, 19.34]                                     | [17.34, 19.34]             |
| 61.5     |                             | [17.67, 24.34]                                   | [10.34, 18.67]                                  | [17.67, 19]                |
| 62       |                             | [18, 24.67]                                      | [10.34, 18.67]                                  | [18, 18.67]                |
| 62.5     |                             | [18.34, 25]                                      | [10, 18.34]                                     | 18.34                      |
| 63       |                             | [18.67, 25.34]                                   | [9.67, 18]                                      | ϕ                          |

The same can be observed graphically in continuous sense as shown below in Figure 1. In this graphical representation mean score has been taken along the x-axis where as standard deviation of the scores along y-axis.
Referring to Figure 1, the mean belonging to \((51.25, 62.5)\) and standard deviation belonging to \((14.34, 21.67)\) covers the marks range \(25 \leq m - 1.5(\sigma) \leq 35\) and \(77.5 \leq m + 1.5(\sigma) \leq 90\). Broadly, with mean lying in the interval \([55, 60]\) and standard deviation lying in the interval \((16.67, 20)\) is a most preferred and acceptable range, as any number combination in the specified intervals will serve the purpose, which is highlighted through a rectangular region in figure 1. However for the mean in \((51.25, 55) \cup (60, 62.5)\) and standard deviation in \((14.34, 16.67) \cup (20, 21.67)\) we refer Table 2 for identifying appropriate ranges for the parameters.

6. Conclusion

Given the evaluation and assessment methods/processes being followed and practicing in current scenario at various levels, it is need of an hour to have an optimal solution or mechanism to address the limitations or controversies associated with them. Thus, the study undertaken conclude that if suggested mechanism provided in our study be adopted, it may quite possible that some of the controversies related to existing evaluation or assessment practices may be addressed to some good extent, if not fully. Further the inference ranges for mean and standard deviation coming out of our study, if it is known to the faculty members well before they start academic delivery, will certainly help them in handling the course and evaluation process effectively in a continuous assessment system. However, the undertaken study also concluded that mechanism proposed in our study may provide some basis for further study depending upon the formula adopted by the different educational institutions, with the use of computation techniques and Information Technology tools to make the things easy and more relevant.

7. References

[1] Ann-Louise de Boer, Pieter H. du Toit, M. Detken Scheepers and Theo J.D. Bothma. 2013 5-
Learning material that makes a difference, Chandos Learning and Teaching Series, 245-272.

[2] Jeffrey Schinske, Kimberly Tanner. 2014 Teaching More by Grading Less (or Differently),
CBE—Life Sciences Education 13 159–166.

[3] EszterCzibor, SanderOnderstal, RandolphSloof, C. Mirjamvan Praag. 2020 Does relative grading help male students? Evidence from a field experiment in the classroom, Economics of
**Education Review**, 75 101953 https://doi.org/10.1016/j.econedurev.2019.101953

[4] Eachempati Prashanti, Komattil Ramnarayan. 2019 Ten maxims of formative assessment, *Adv Physiol Educ* 43 99 –102.

[5] Jonathan D. Kibble. 2017 Best practices in summative assessment, *Adv Physiol Educ* 41 110–119 doi:10.1152/advan.00116.2016.

[6] Lefebvre M, 2018 Applied Probability and Statistics, *Springer*.

[7] Malini Reddy Y., Heidi Andrade. 2010 A review of rubric use in higher education, *Assessment & Evaluation in Higher Education* 35(4) 435–448.

[8] Azmat G, and Iriberri N. 2010 The importance of relative performance feedback information: Evidence from a natural experiment using high school students, *J Pub Econ*. 94(7) 435–52. https://doi.org/10.1016/j.jpubeco.2010.04.001.

[9] Erin B.D. Clabough, Seth W. Clabough. 2016 Using Rubrics as a Scientific Writing Instructional Method in Early Stage Undergraduate Neuroscience Study, J Undergrad Neurosci Educ. 15(1) A85–A93.

[10] Andreoni J, Brownback A. 2017 All pay auctions and group size: Grading on a curve and other applications, *J Econ Behav Organ.* 137 361–373. https://doi.org/10.1016/j.jebo.2017.03.017.

[11] Abadie A, Athey S, Imbens G W, and Wooldridge J. 2017 When should you adjust standard errors for clustering?, *National Bureau of Economic Research Working, Paper* 24003.

[12] Brownback A. 2018 A classroom experiment on effort allocation under relative grading. *Econ Educ Rev* 62 113–128. https://doi.org/10.1016/j.econedurev.2017.11.005.

[13] Campos-Merceade, P., and E. Wengströ’m. 2019 Threshold incentives and academic performance. Job market paper. https://www.dropbox.com/s/ejnc8uzz6y30dio/Campos-Merceade_JM.pdf?dl=0. pp 1-70.

[14] Shinya Kajitani, Keiichi Morimoto, Shiba Suzuki. 2020 Information feedback in relative grading: Evidence from a field experiment, *PLoS ONE* 15(4): e0231548. https://doi.org/10.1371/journal.pone.0231548.

[15] Kohn A. 1999 Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A’s, Praise, and Other Bribes, *Houghton Mifflin Harcourt*, New York.

[16] Jascik S, 2014 Imagining College without Grades, www.insidehighered.com/news/2009/01/22/grades.

[17] Levitt S D, List J A, Neckermann S, Sadoff S. 2016 The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance. *Am Econ J Econ Policy.* 8(4) 183–219. https://doi.org/10.1257/pol.20130358.

[18] Andy Brownback. 2018 A classroom experiment on effort allocation under relative grading, *Economics of Education Review* 62 113-128.

[19] Dirk Witteveen, Paul Attewell. 2020 The STEM grading penalty: An alternative to the “leaky pipeline” hypothesis, *Science Education*, 104(4) 714-735.

[20] Kajitani S, Morimoto K, Suzuki S. 2020. Information feedback in relative grading: Evidence from a field experiment. *PLoS ONE* 15(4): e0231548. https://doi.org/10.1371/journal.pone.0231548.

[21] Mahendra Parihar, Nagaraju Dasari. 2019 “Using Intelligent Communication System (ICT) to Solve a Problem of Enroute Confirmation of Wait Listed Tickets in Indian Railways”, *IEEE transactions, 2019 2nd International Conference on Intelligent Communication and Computation Techniques*, 288-291.

**Acknowledgement**

Authors of the paper would like to express their gratitude for the necessary guidance received from Manipal University Jaipur, Mukesh Patel School of Technology Management, SVKM’s NMIMS (Deemed to be University), Mumbai and Narasaraopeta Engineering College, Guntur.