Distribution of Test Problems to Ensure Ideal Distribution

Control Number #905

AoCMM 2015

Pratham Gupta, Angela Yi, Kevin Wu, Richard Liu

September 11, 2015 to September 13, 2015

b. Summary

Interpretation of contest problem: How should a professor distribute the difficulty of problems on a test to ensure that a group of students, with varying abilities, will form an ideal distribution? The contest problem's usage of the word professor implies that we are referring to a college or university setting. A group of students constitutes an average class of students in a university (>20). Students with varying abilities implies a differing level of intelligence, knowledge, and reasoning ability. It is up to individual teams to define an ideal distribution.

Methods used: Microsoft Excel was used to model Gaussian normal distribution of problems of different difficulty. The Gaussian model was also our definition of an ideal distribution. By taking a sample of the group of students' intelligence through IQ scores, which were always normalized, we adjusted the mean of the new ideal distribution (input) to show the curve of grades (output).

Conclusions: The objective of our model was achieved. Professors can simply adjust the mean of the Gaussian curve (ideal distribution) of grades to determine the distribution of test questions with different difficulty. The Gaussian curve for the distribution of grades would be skewed right and left as the mean increased and decreased. The distribution of the different difficulties of problems would shift identically with the distribution for grades. One standard deviation on the normalization of grades was equivalent to one level of difficulty. If the Gaussian curve for grades shifts right as the mean increased, the curve for the distribution of difficulties of problems would shift right in the same manner. Please note that increasing the mean requires the increase of easier test questions in order to make a higher score more attainable.

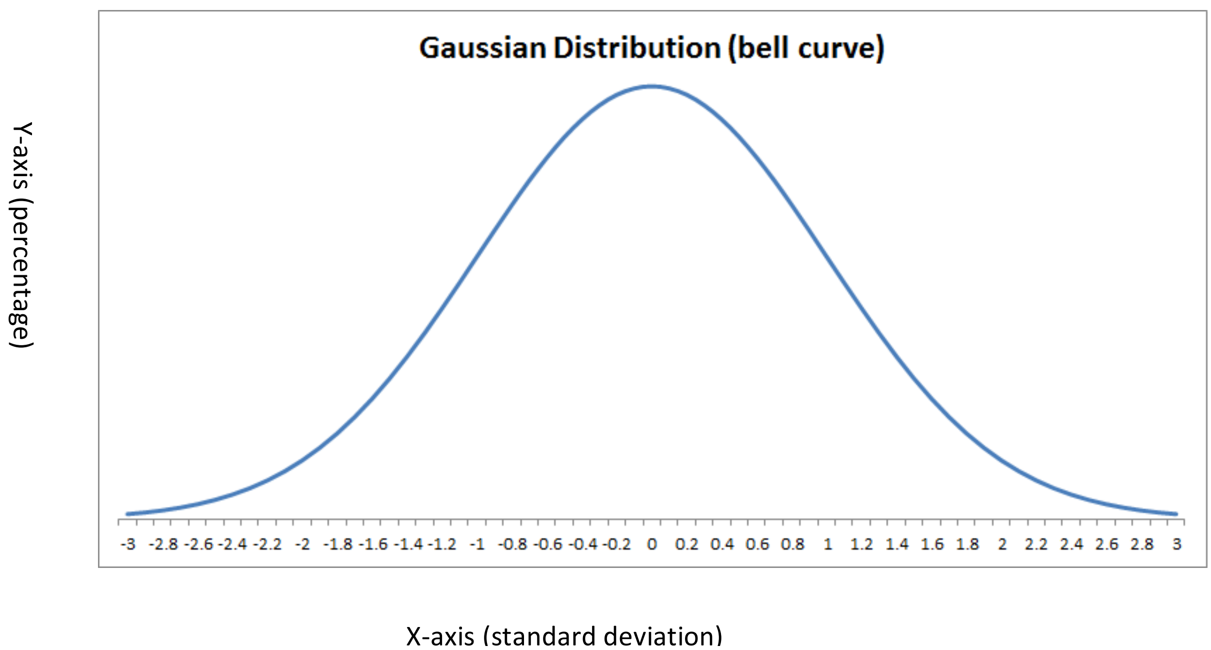
c. Table of Contents

d. Introduction

How should a professor distribute the difficulty of problems on a test to ensure that a group of students, with varying abilities, will form an ideal distribution?

Currently, professors will create tests with a set of problems that they have created without regards to the varying and specific ability levels of each student. After the tests are graded, and professors will use several methods to curve the scores. The first is the Gaussian distribution, or the bell curve. There spread of data is distributed into a bell curve where the majority of data is concentrated near the mean.

e. Model



i. Introduction of the model

There must be an objective method of measuring the varying abilities of new students, which professors have no prior knowledge of. We believe that IQ (Intelligence Quotient) testing performed by professionals is the best way to show students' abilities. With the raw data from the testing, professors can develop a standard distribution of IQ scores in a class using the Gaussian normal distribution model. The students in a class will thus be divided into several levels of relative intelligence and reasoning ability. Professors will write future tests based on this data.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

This first figure on the left is the equation for standard deviation. Given a set of data size □, first calculate the mean of the set. In this equation, μrepresents the mean. (Usually □ represents the mean). For every □□, the mean is subtracted from it and then squared. Take all □ of these modified values and sum them up. Finally, divide the sum by □ and square root it. This results in the standard deviation.

Another form of the standard deviation divides the sum by □ − *1* instead of □, but that involves a sample and population. In this case, we are using standard deviation to find an ideal distribution, which requires dividing by □ instead of □ − *1*.

Normally, the ideal distribution should have 99.8% of the data/values within the interval [□ − 3□, □ + 3□] , where □ is the mean and $\sigma$is standard deviation. The values □ − 3□and□ + 3□are included in the interval.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

This figure on the left is the function of the Gaussian distribution. Given a data set and its standard deviation, the probability density can be calculated using this equation. The probability density is the probability that a value □is selected out of the data.

ii. Assumptions and Justifications

In order to further strengthen our model, we made the following assumptions:

Assumption: The maximum possible score on a test is 100% and the minimum is 0%

Justification: We used the standard parameters for a modern college and university exam. Extra credit and negative deductions are anomalies, and thus we did not factor them into the model

Assumption: Professors and research/education professionals can design questions with specific levels of difficulty

Justification: Standardized tests across the board have always contained problems with quantitative levels of difficulty. It will not be difficult for professors to develop tests in a similar manner, single-handedly or in consultation with education professionals.

Assumption: Gaussian/normal distribution can be reached

Justification: For the Gaussian/normal distribution model to properly function, students must have different levels of ability. The problem already clearly establishes that students will have a varying range of abilities.

Assumption: IQ tests are the best of way of measuring the varying abilities of students.

Justification: Longtime education journalist Michael Balter explains that students who "score higher on IQ will, on average, go on to do better in...academic achievement." As *The Atlantic* shows, it is generally accepted that an IQ test reveals a person's native intelligence. While often prone to fluctuations, it is a much stabler value in comparison to all other standardized tests. State standardized testing and college entrance exams such as the SAT/ACT test knowledge and critical thinkings skills at particular points in time, and are not representative of a person's innate ability.

iii. Clear explanation of solution

We will ask all students in a class to take an initial IQ (intelligence quotient) test. IQ test scores are already in a normalized or Gaussian distribution. The mean of the normal distribution will then be shifted to a desired test mean in the "ideal distribution," which remains normalized (though may be skewed) with the same standard deviation. For the next test, the ideal mean of the normal distribution will be determined by a professor. The common Gaussian distribution with a width of 1 standard deviation is commonly divided into 8 percentage categories, and thus 8 levels of relative intelligence and ability. To design a test that reaches this ideal distribution, professors must create 8 levels of questions at different difficulty in accordance with relative intelligence levels. These 8 levels must also be distributed by percentage corresponding to the normal distribution. For example, if a professor wants the students in interval $\underline{\square} + 1\square$ to score a B+ on an upcoming exam, he would make the questions in interval $\underline{\square} + 1\square$ at an easier than average difficulty. In order for more students to achieve a higher grade, the questions must be easier (thus more easily attainable by more students).

iv. Sensitivity Analysis

Figure A:

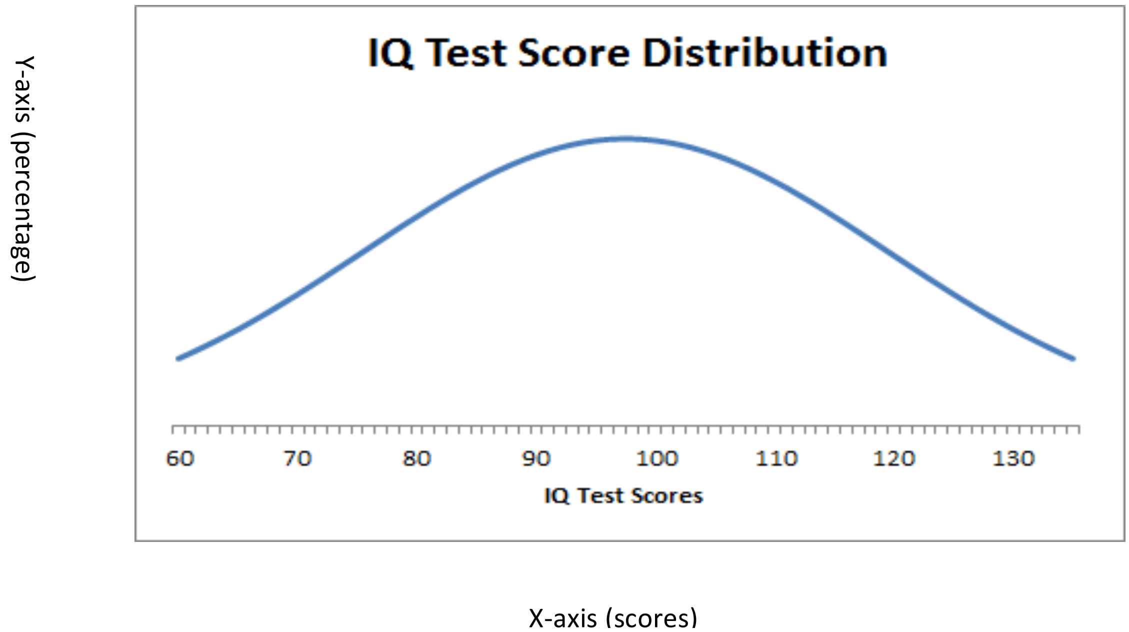


Figure A represents a sample Gaussian Distribution of the scores of the people taking the IQ Tests.

Figure B:

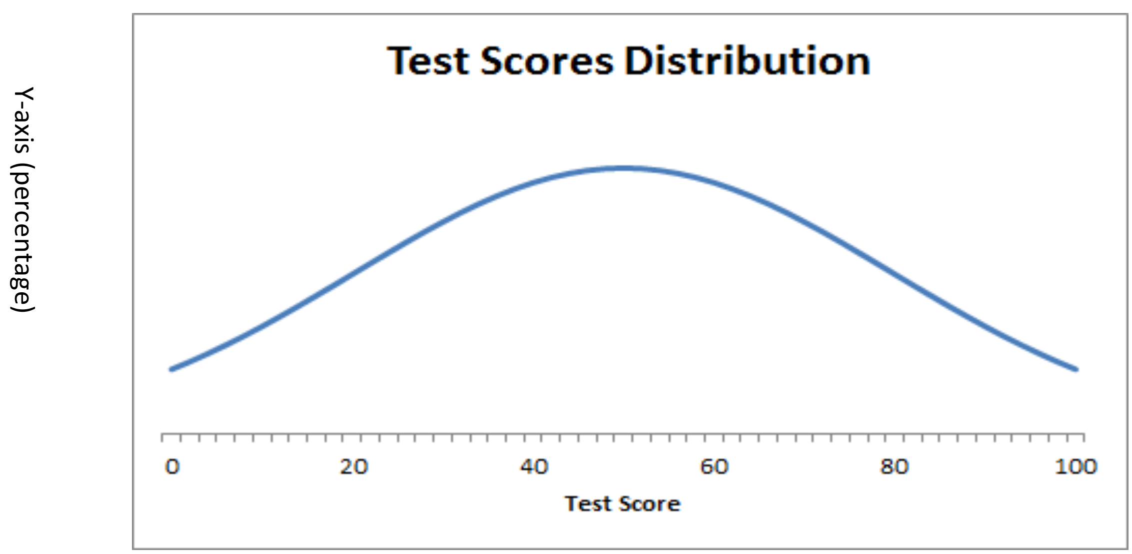


Figure C:

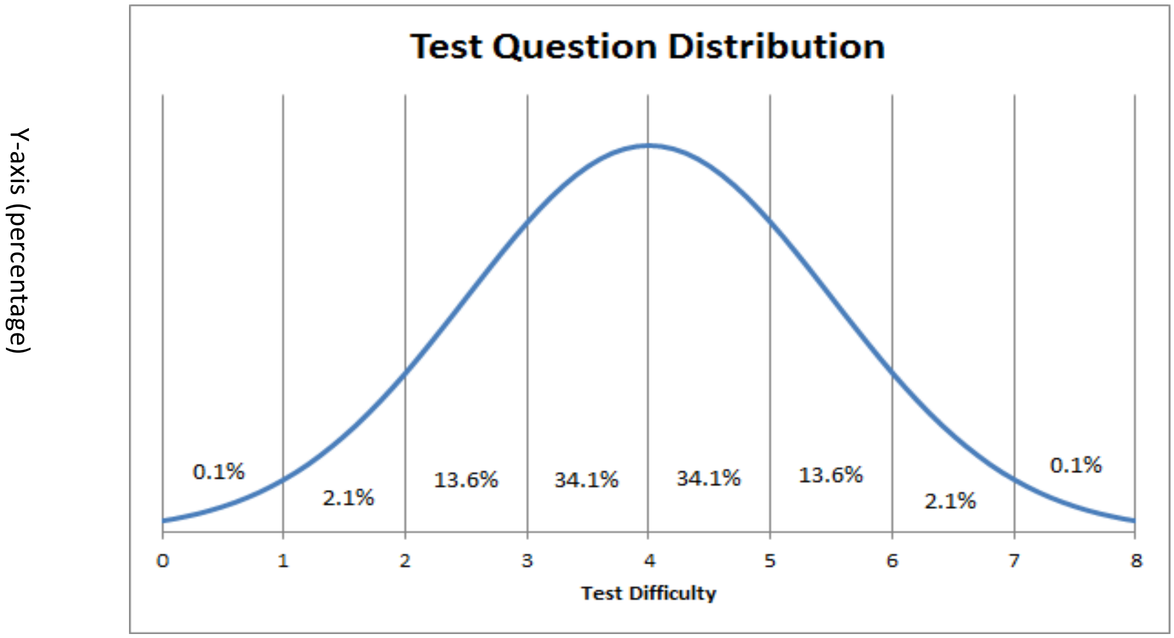


Figure B represents one ideal distribution of the test scores. The scores are distributed in a standard Gaussian/normal curve with the mean at 50. 1 standard deviation is set at 12.5 points. The corresponding distribution of test questions is in Figure C. The test questions are designated on a scale of 0-7 for difficulty. Questions rated at 0 are the hardest possible, and ones rated at 7 are the easiest possible. 1 standard deviation is 1 level of difficulty.

Figure D:

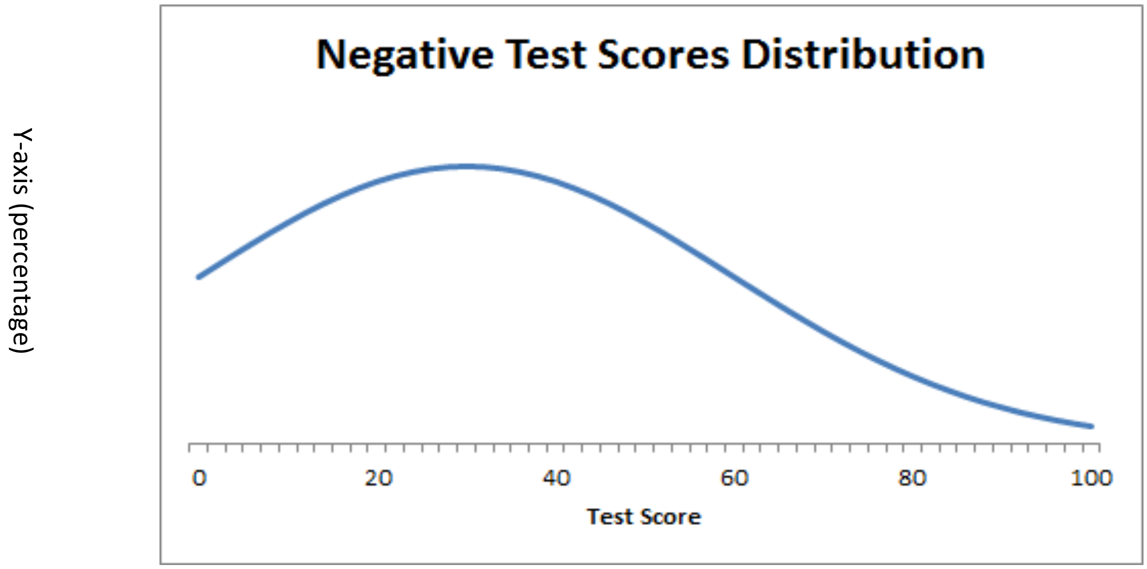
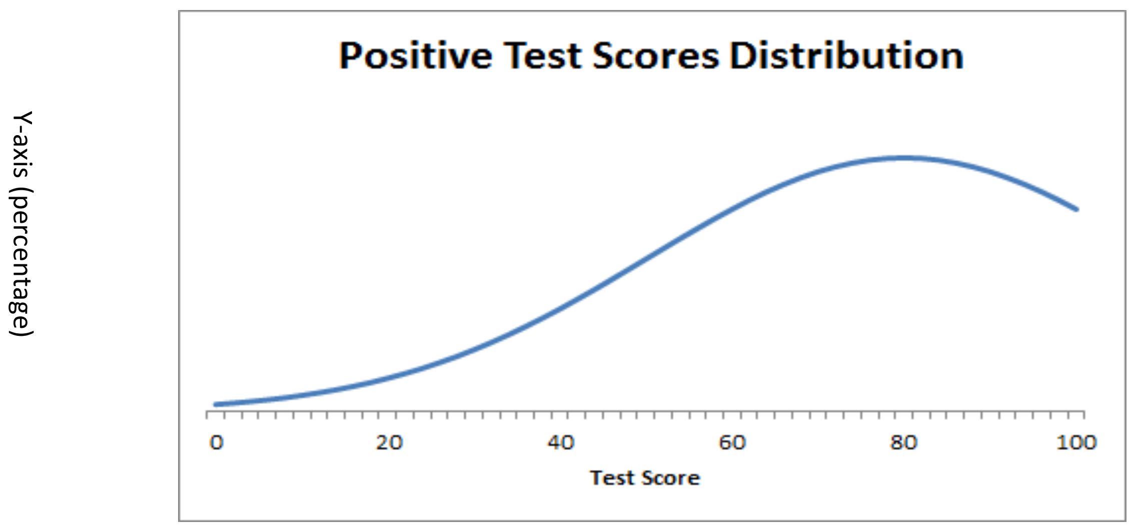


Figure E:



In Figure D at the top of the page, the professor has shifted the mean to 30. The graph of the Gaussian distribution becomes a Skewed Right Graph. Most of the students will end up with a lower grade. In this example, the professor will have to create more difficult questions in direct accordance with the distribution. In Figure E, the professor has shifted the mean to 90. The graph of the Gaussian distribution becomes a Skewed Left Graph. Most of the students will end up with a higher grade. In this example, the professor will have to create more easy questions in corresponding with the percentages in the curve.

v. Strengths and Weaknesses

Strengths

- The Gaussian distribution is almost always applicable. It is probably the most widely known distribution method and is used for most of all distributions. Intelligence quotient (IQ) scores are always normally distributed, and test scores most often follow the curve as well.
- The Gaussian distribution is also very easy to work with mathematically. It can be manipulated using algebra (since it has a formula) much more easily than other types of distributions. It is possible to derive results that can be easily applied.

Weaknesses

- By using the Gaussian distribution, you would need to assume that there are not many extreme scores that will distort the mean. Distorting the mean would make the graph an inaccurate representation of the distribution of scores and questions.
- In a real life situation, the mean and thus the distribution of scores and questions may change slightly from test to test. The first curve may be skewed left, while the next mean/graph may be a few points off.
- The model assumes that standard deviation is constant. In cases where the standard deviation does change, more error will appear in both extremes of the Gaussian curve (i.e. grades A and F).

f. Conclusion

After IQ scores were gathered from a class (which are always normalized), a professor will shift the mean of the scores to any desired value to generate the ideal Gaussian distribution for scores on a test. For professors to determine the ideal distribution of test questions of different difficulty, they can simply shift the mean of the Gaussian distribution of scores. This is because the distribution for ideal test scores is identical to the distribution for test questions of different difficulty. Any shift in the curve for scores/grades would be reflected in the curve for difficulty as well. One standard deviation on the distribution of grade would be equivalent to one level of difficulty. Increasing the mean requires the increase of easier test questions in order to make a higher score more attainable, and vice versa. Potential weaknesses include the ability of outliers and extremes to distort the distribution. Also, the mean/distribution of scores may shift slightly from test to test. A change in standard deviation would also affect the extremes of the Gaussian curve. However, the model is still very strong. The Gaussian distribution is widely used and applicable in the real world. It is also very simple to modify and shift. Any outliers or slight shift as described in the weaknesses will not significantly impact the overall validity of the model.

g. Citations

Balter, Michael. "What Does IQ Really Measure?" *Advancing Science, Serving Society*. AAAS, 25 Apr. 2011. Web. 12 Sept. 2015.

Brook, Stephen E. "Descriptive Statistics and Psychological Testing."Descriptive Statistics and Psychological Testing (2011): 1-7. Web. 12 Sept. 2015.

De La Jara, Rodrigo. "IQ Percentile and Rarity Chart." IQ Comparison Site. IQ Comparison Site, n.d. Web. 12 Sept. 2015.

Lund, Adam, and Mark Lund. "Measures of Central Tendency." Laerd Statistics. Laerd Statistics, 2013. Web. 12 Sept. 2015.

"The Mean and Standard Deviation." Cosmic Ray Observatory Project. University of Nebraska-Lincoln: Physics & Astronomy, 2015. Web. 12 Sept. 2015.

Shenk, David. "The Truth About IQ." *The Atlantic*. Atlantic Media Company, 28 July 2009. Web. 12 Sept. 2015.

"Standard Deviation and Variance." Math Is Fun. Math Is Fun, 2014. Web. 12 Sept. 2015.

Yount, Rick. "The Bell Curve and Assigning Grades." Dr. Rick Yount - Teaching That Helps Learners Grow in the Lord. Dr. Rick Yount, 30 June 2011. Web. 12 Sept. 2015.