



A Visual and Engaging Approach to Teaching and Learning the Normal Distribution

Dr. Daniel Raviv, Florida Atlantic University

Dr. Raviv is a Professor of Computer & Electrical Engineering and Computer Science at Florida Atlantic University. In December 2009 he was named Assistant Provost for Innovation and Entrepreneurship.

With more than 25 years of combined experience in the high-tech industry, government and academia Dr. Raviv developed fundamentally different approaches to "out-of-the-box" thinking and a breakthrough methodology known as "Eight Keys to Innovation." He has been sharing his contributions with professionals in businesses, academia and institutes nationally and internationally. Most recently he was a visiting professor at the University of Maryland (at Mtech, Maryland Technology Enterprise Institute) and at Johns Hopkins University (at the Center for Leadership Education) where he researched and delivered processes for creative & innovative problem solving.

For his unique contributions he received the prestigious Distinguished Teacher of the Year Award, the Faculty Talon Award, the University Researcher of the Year AEA Abacus Award, and the President's Leadership Award. Dr. Raviv has published in the areas of vision-based driverless cars, green innovation, and innovative thinking. He is a co-holder of a Guinness World Record. His new book is titled: "Everyone Loves Speed Bumps, Don't You? A Guide to Innovative Thinking."

Dr. Daniel Raviv received his Ph.D. degree from Case Western Reserve University in 1987 and M.Sc. and B.Sc. degrees from the Technion, Israel Institute of Technology in 1982 and 1980, respectively.

Mr. Daniel Ryan Barb, Florida Atlantic University

Daniel Barb is an undergraduate student studying Mechanical Engineering at Florida Atlantic University. He spent six years in the United States Navy working in a nuclear power plant aboard a fast attack submarine.

A Visual and Engaging Approach to Teaching and Learning the Normal Distribution

Daniel Raviv and Daniel Barb
College of Engineering and Computer Science
Florida Atlantic University
Emails: ravivd@fau.edu, dbarb@fau.edu

Abstract

Probability and Statistics is a pervasive topic in many careers. Engineers, medical professionals, and other specialists use Statistics to analyze data and determine courses of action. People utilize the basic concepts of probability to make decisions in their everyday lives. Despite this, many students struggle with Probability and Statistics in classroom settings. The equations and mathematics can be overwhelming or frustrating, partially because lessons are not visual or intuitive enough and are often missing the connection to daily, real-world experiences.

This paper focuses on a visual, intuitive, example-based, and engaging approach to teaching a specific Probability and Statistics concept, namely the normal distribution. The point of this approach is to provide students with easy to understand examples that translate bookwork to real life and help in comprehension of the material. This more intuitive understanding allows them to be more successful when introduced to equations and calculations associated with the topic. To teach this concept, the paper shares supplemental materials that can be introduced during class instruction prior to delving into mathematical details. The paper introduces this concept using (1) Examples, such as a Galton Board and flipping coins, (2) Visualizing basic concepts and some key concepts, (3) Real-life, experience-based examples such as height distribution, (4) A puzzle involving a multiple-choice exam, and (5) An in-class experiment of rolling a die.

It should be noted that this paper is a work in progress. In addition, this method of teaching is meant to be supplemental and not to replace existing textbooks or other teaching and learning methodologies. The work in this paper has been presented to 21 students in a Probability and Statistics classroom setting. Following the presentation, it has been assessed and received very positive feedback. This visual, intuitive, and engaging approach to teaching and learning has also been tested in the past for topics in Statics (explaining center of gravity), Calculus (explaining integration and explaining derivation by chain, product, and quotient rules), Differential Equations, Control Systems, Digital Signal Processing, Newton's Laws of Motion, and Computer Algorithms. In all of these cases, students commended the approach and found it to be very effective for learning.

I. Introduction

Most Statistics textbooks are loaded with mathematical formulas and explanations with little focus on conceptual understanding. This is of course a great approach to mathematically introduce the material since it is written in a precise manner, but at the same time students may become frustrated as they do not intuitively grab some of concepts, and miss the “aha” moment. Therefore, instructors must modify their teaching approach to first introduce the concept in a clear, easy to understand way, and only later delve into the textbook math.

Many books have been written on the subject of Probability and Statistics. Most textbooks, for example [1], generally jump straight into the mathematics. Other books, though usually not textbooks, try to introduce Statistics in a visual way. For example, [2] uses comics as an approach to teaching the subject, which helps visual students understand the basic ideas, and [3] gives several visual examples from real life while clearly explaining mathematics to general audiences. [4] and [5] also focus on explaining the topic to a more general audience, trying to make Statistics fun for the layman. A few professors have understood the need for explaining Statistics differently, and came up with some alternative approaches: for example, [6] includes many tricks and hands-on activities for students. Books like [7] explain the concepts in a textbook-like manner. Others, for example [8] and [9], try to build up the reader’s intuition through real life examples.

These books are useful supplements, but if students miss concepts it is often hard to readjust their thinking and alleviate their frustrations. The result of this is that students often struggle with, or completely miss, a basic understanding of concepts, and don’t get the “Aha!” moment that makes things click.

The focus of this paper is to provide a set of supplementary material that can help instructors to teach, and students to comprehend, a specific Probability and Statistics concept, namely the normal distribution. By introducing this concept in visual, intuitive, and engaging ways, the material becomes more meaningful for the student, particularly when they can relate it to real-life.

This paper starts with intuitive, experience-based examples to introduce the concept, such as the Galton Board and coin flips, followed by introducing key concepts, such as an answer to the simple question, “What does a bell curve mean?” It then provides real-life examples of the normal distribution, such as height distribution of classmates. Finally, it ends with a puzzle and an in-class experiment that can engage students in a classroom setting.

The work in this paper has been presented to 21 students in a Probability and Statistics classroom setting. Following the presentation, its effectiveness was assessed and received very positive feedback, which is shown in detail in Appendix A.

II. Some Intuitive, Experience-Based Examples

This section describes three real-life examples that can be used to introduce the basic idea behind the normal distribution. These are the Galton Board, flipping coins, and driving in a city.

A. The Galton Board

The Galton Board (Figure 1), also known as a bean machine or quincunx, distributes marbles between vertical compartments. If you were to drop a marble into the Galton Board like the one below, the marble could end up in any of several vertical columns. The specific outcome for any single marble cannot be accurately predicted. As more marbles are dropped into the board from the top, they will end up in different columns with varying frequencies: most will end up around the center, and less will end up in the far right and left columns (outliers). Given that many marbles are dropped, the stacks will assume a bell shape. A note: it can be shown that this bell shape approximates the normal distribution. Why is this so?

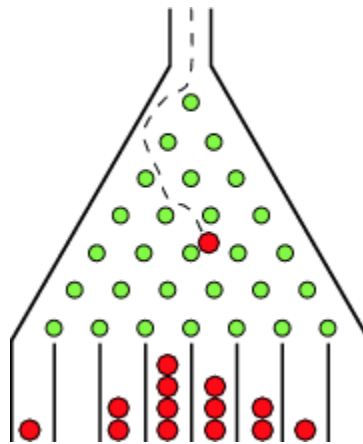


Figure 1: Galton Board

A marble entering the Galton board initially has only one path to follow (Figure 2), and the marble drops directly on top of the top peg. From there it can continue in one of 2 equally likely ways: to the left or to the right. Each path has a 50% chance of occurrence (Figure 3).

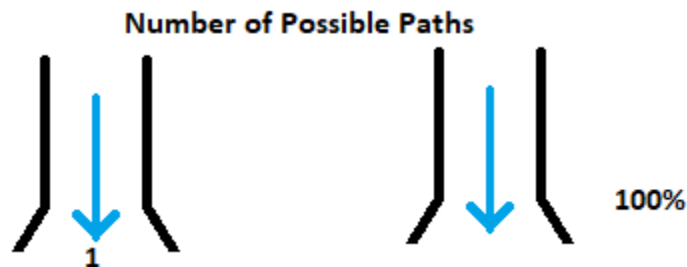


Figure 2: Galton Board with only one possible path

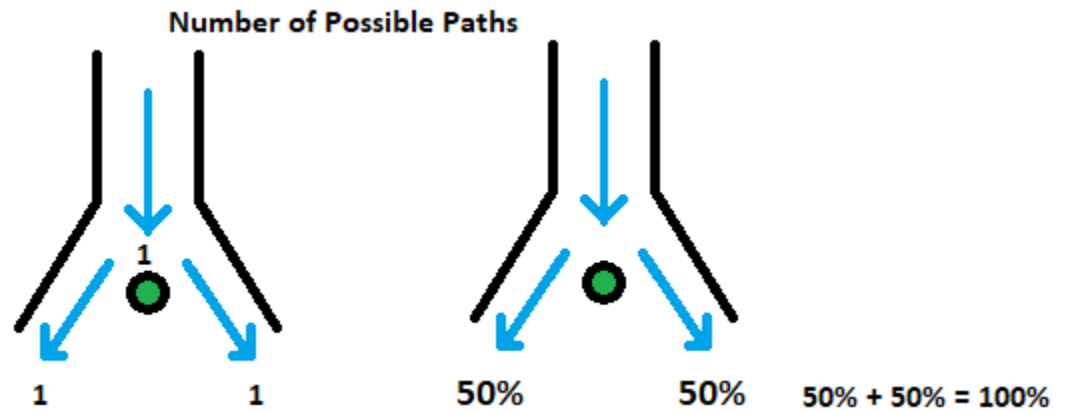


Figure 3: Galton Board with only the top peg considered

We now consider the second row. Again, each marble that hits a peg can continue either to the left or the right (where each outcome is equally likely). Given that there is a 50% chance for a marble to get to any given peg on the second row, there is therefore a 25% chance for a marble to fall to either side of the next level peg. We therefore have four possible outcomes, each with a 25% chance of occurring, but two of these outcomes (right then left, and left then right) take the marble to the same place! Therefore, the outside destinations have one possible chance of occurring whereas the middle destination has two possible chances of occurring. Thus, when we consider the marble's destination and not the path taken, a single marble has a 25% chance to be on the left, a 25% chance to be on the right, and a 50% chance to be in the middle, as shown in Figure 4.

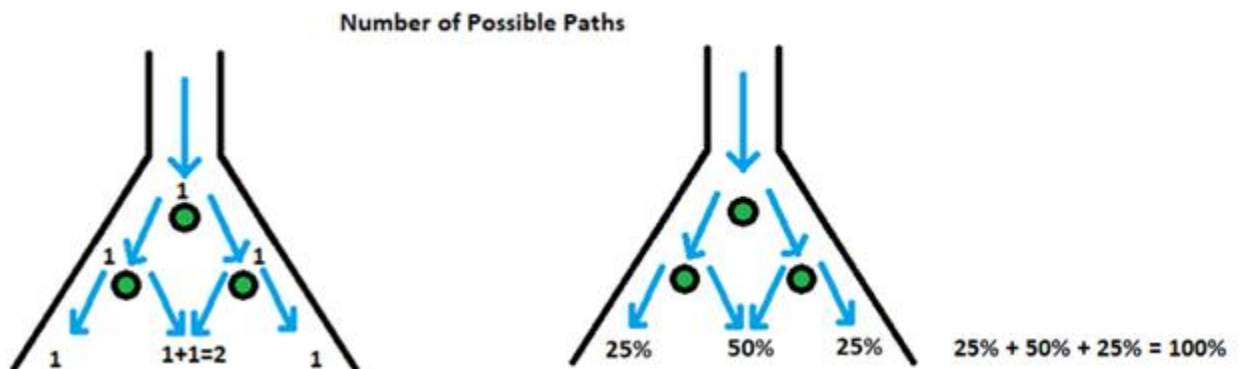


Figure 4: Galton Board with the first two rows of pegs considered.

The figure on the left shows the number of possibilities for reaching a given location on the board, and the figure on the right shows the probability of each outcome

When asking the question: “In how many different ways can a marble get to a certain level on the Galton Board?” we find that each level has twice as many possible paths as the previous level. That is, there are twice as many ways to get to a given level than to get to the previous level. This is clear in Figure 5, where the number of paths are 1, 2, 4, and 8. This is because at any peg a marble can continue to the left or the right. Note that at any level the marble

is 100% sure (a “sure event”) to reach the next level of the Galton Board at some position (Figure 6).

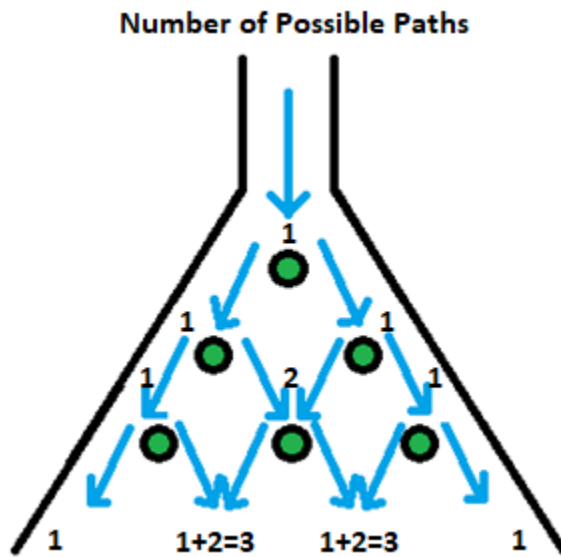


Figure 5: Possible paths at the first few levels

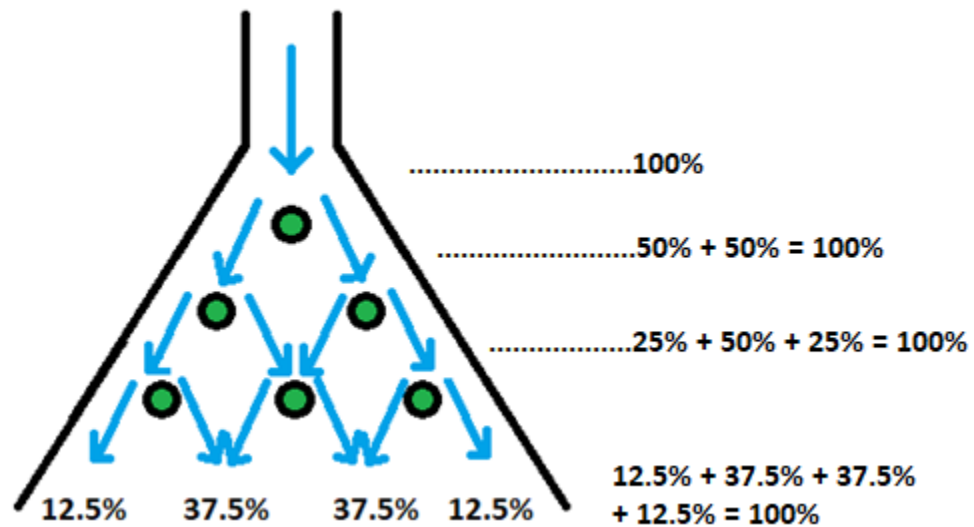


Figure 6: Galton Board with the first three rows of pegs considered

As a marble makes its way to lower levels, it encounters rows with more pegs, and the probability of the marble ending up at (or near) the center column becomes higher compared to the extreme left and right columns, as shown in Figure 7.

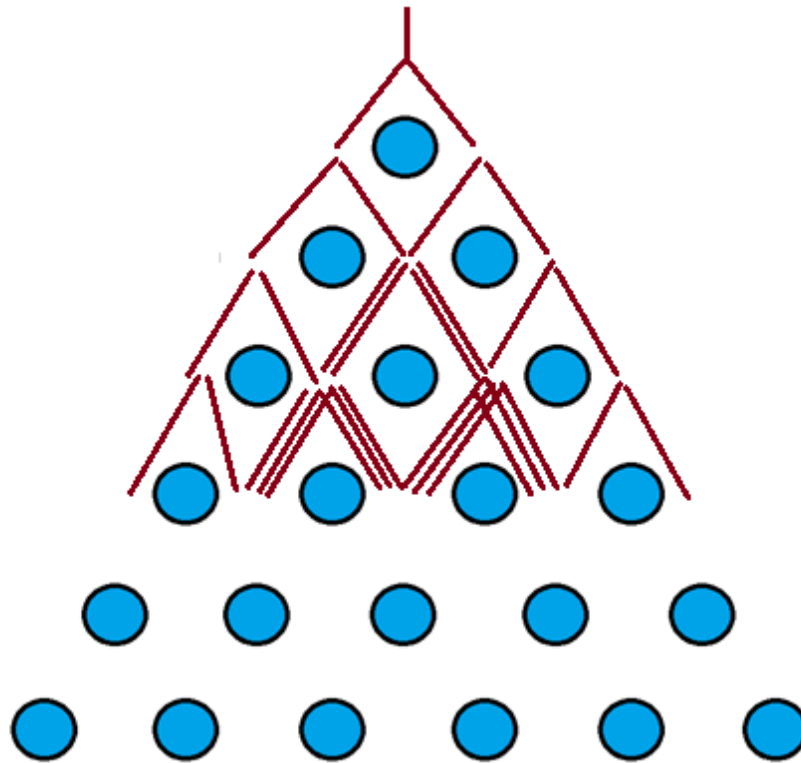


Figure 7: Possible paths for the first few rows of the Galton Board

If we count the number of possibilities for getting to a specific point, we get a chart as shown in Figure 8. This theoretically infinitely long chart is known as Pascal's Triangle, and is discussed in Appendix B.

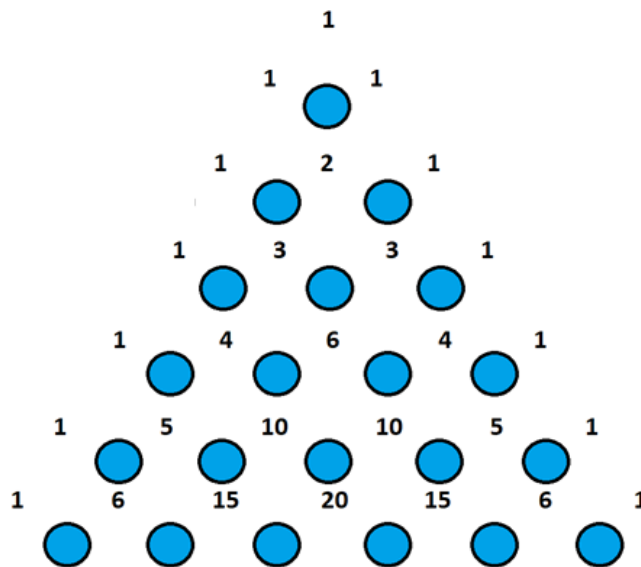


Figure 8: The number of possibilities for each row of pegs

Note that each number is the sum of the two numbers above it. For example, (referring to figure 9) $4+6=10$. This means that if there are 4 possible paths for a marble to arrive at the top left peg and 6 possible paths to arrive at the top right peg, then the peg below and between these two pegs has 10 possible paths in which a marble can arrive.

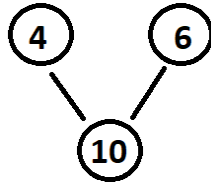


Figure 9: Each number in Pascal's Triangle is the sum of the numbers above it.

As we continue adding more rows of pegs, we can plot the probabilities of a marble's arrival at a given point on the Galton Board (Figure 10, left). As more rows are considered, the probability chart approaches a bell shape (Figure 10). As mentioned earlier, this bell shape, for large enough experiments, approaches the normal distribution.

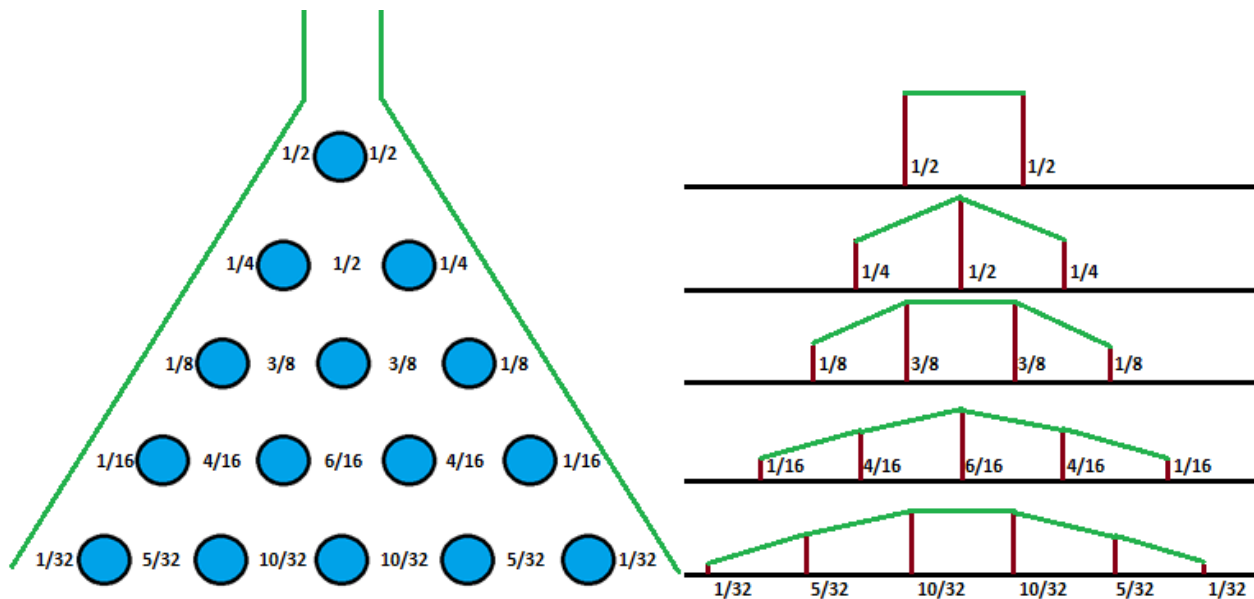


Figure 10: As more rows of pegs are added, the probability pattern assumes a bell shape

For example, let us consider the row of the Galton Board with 9 pegs (the bottom row). The probabilities are shown in Figure 11.

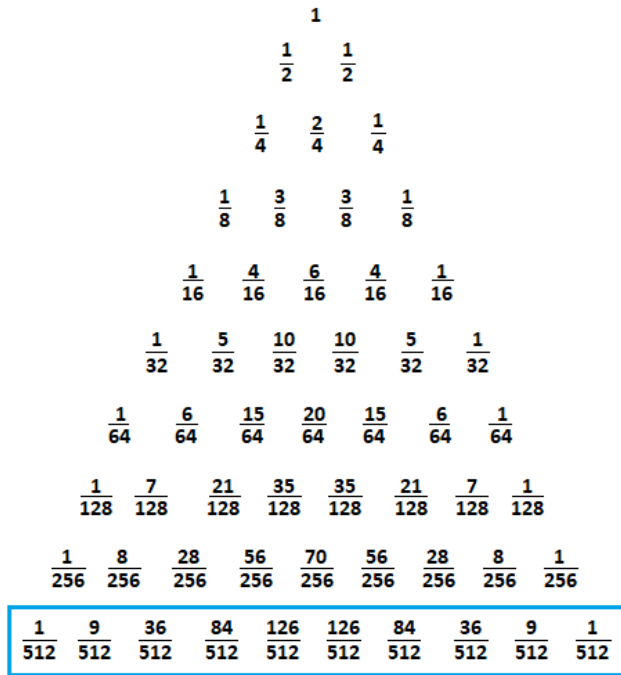


Figure 11: Probabilities shown for the first 9 rows of pegs in a Galton Board.

If we plot the probabilities for the bottom row, we get the graph shown in Figure 12. Note that the sum of all probabilities for this row, like all other rows, is equal to 1. That is to say, it is a sure event that a marble will reach that level (or any other).

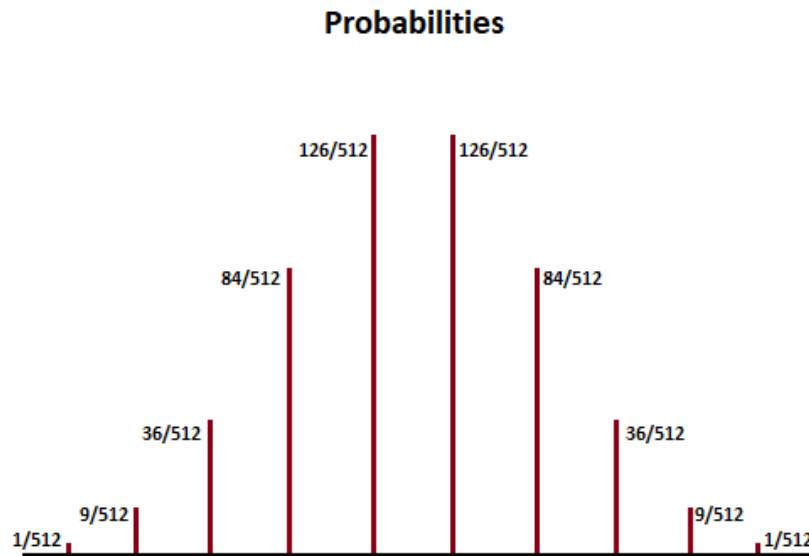


Figure 12: Probabilities shown for the 9th row of pegs in a Galton Board.

If we connect the points on the graph, we get the graph shown in Figure 13. We can use this to approximate a bell-shaped graph, as shown in Figure 14.

Probabilities Connecting with Straight Lines

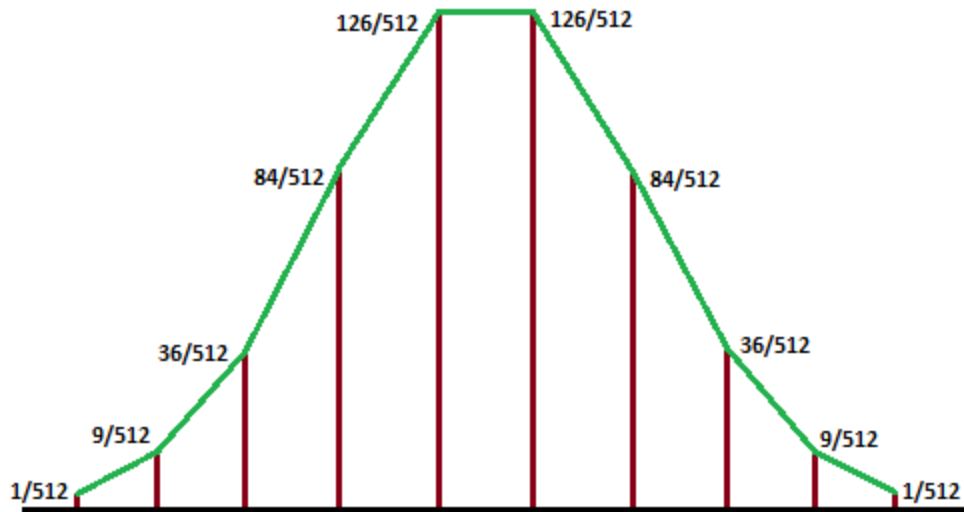


Figure 13: Connecting the previous graph with straight lines from point to point.

Probabilities Approximating with Bell shaped Graph

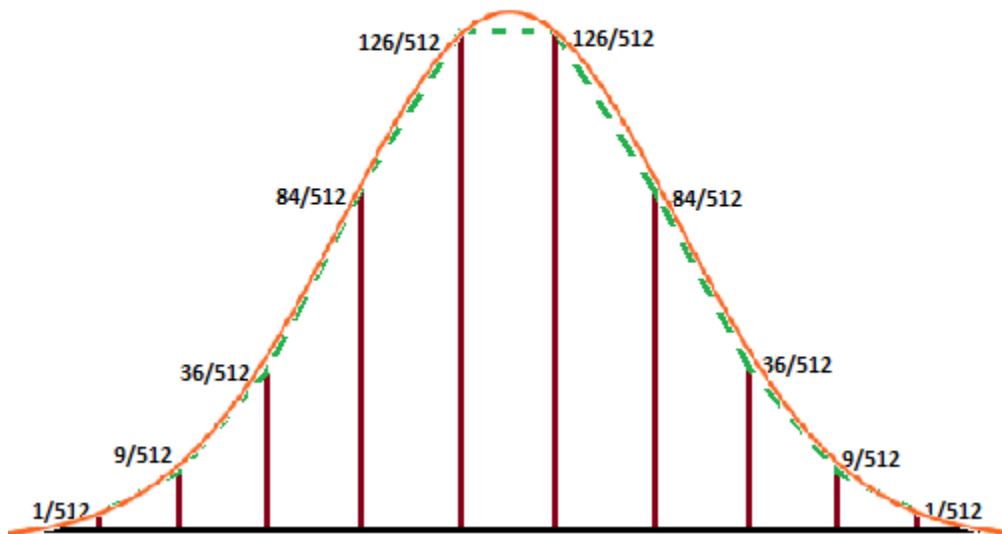


Figure 14: Approximating the probabilities for different points with a continuous curve. This graph forms a bell-shape

Note the difference in graph shapes between Figure 14 and Figure 10. Figure 14 looks much more like a bell curve than the curves in Figure 10. As the number of rows of pegs in a

Galton Board increases, the graph assumes a shape of a bell, as shown in Figures 15 and 16. As mentioned earlier this bell shape graph approximates the normal distribution.

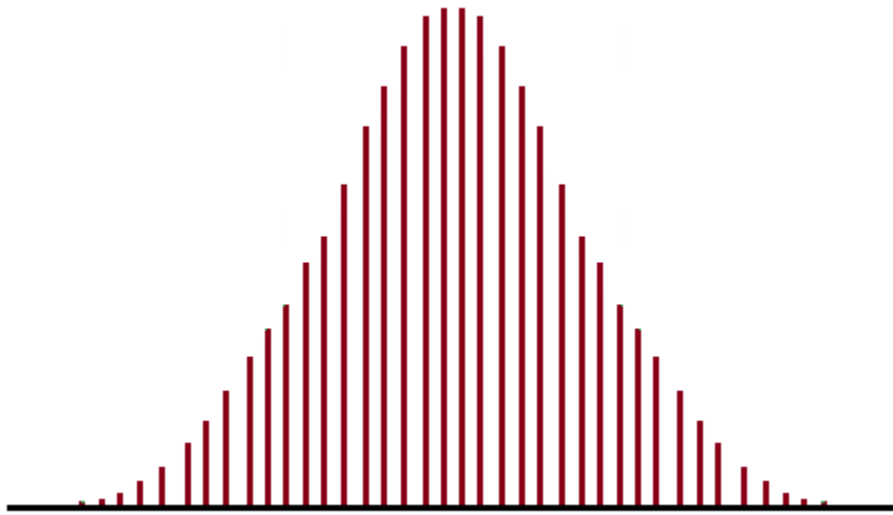


Figure 15: A graph with discrete probability values representing probabilities for the 37th row of a Galton Board

Bell Shape Approximation

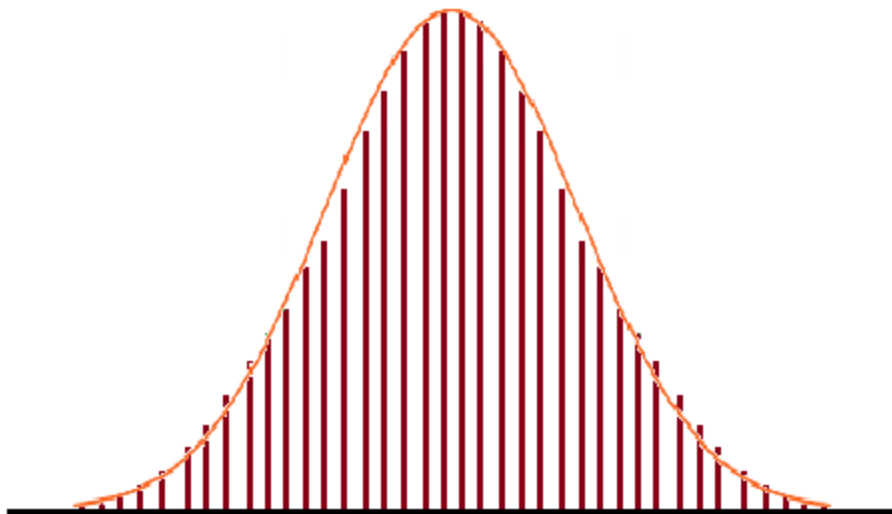


Figure 16: The graph from Figure 15 with an overlaid bell-shaped curve.

B. Coin Flips

Many people use a coin flip to make a decision. Assuming that coins are fair, i.e., there is the same probability ($1/2$) for a “head” as for a “tail.” Since a coin flip only has two possible outcomes, it is known as “binomial.” Figures 17 and 18 show the probabilities for one coin flip.

1 Coin Flip	
Outcome	Probability
H	50%
T	50%

Figure 17: Possible outcomes for a single coin flip.

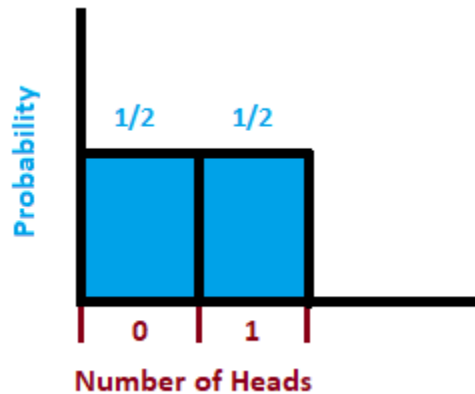


Figure 18: Probabilities of one coin flip

We can analyze the outcomes for two coin flips, like in Figures 19 below. It is important to note that each coin flip is independent of every other coin flip, and thus the outcome of the first flip does not affect the second.

2 Coin Flips		
Outcome	Number of Heads	Probability
T,T	0	25%
T,H	1	25%
H,T	1	25%
H,H	2	25%

Figure 19: Possible outcomes for two coin flips.

Since two of the outcomes (tails followed by heads and heads followed by tails) are both the same, i.e., both result in one heads and one tails (since the order is not important), this outcome has two possibilities for occurrence whereas the other outcomes, i.e., two tails in a row or two heads in a row, have only one possibility for occurrence. Thus, the outcome of “one heads and one tails” is twice as likely as “both tails” or “both heads” as shown in Figure 20.

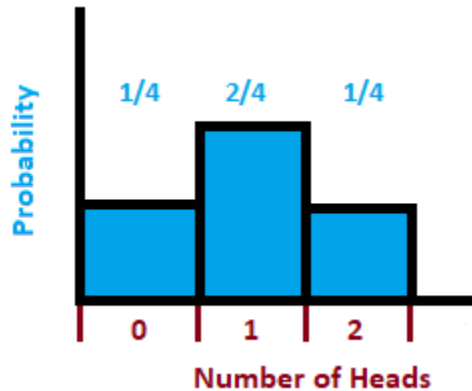


Figure 20: Probabilities for two coin flips

Figure 21 charts the possible outcomes for one coin when flipped three times. Note that there are 8 different possibilities, all with equal probability ($1/8$). If we care only about the number of heads and tails and not the order in which the H's and T's have appeared then we can plot the result (shown in Figure 22).

3 Coin Flips		
Outcome	Number of Heads	Probability
T, T, T	0	12.50%
T, T, H	1	12.50%
T, H, T	1	12.50%
T, H, H	2	12.50%
H, T, T	1	12.50%
H, T, H	2	12.50%
H, H, T	2	12.50%
H, H, H	3	12.50%

Figure 21: All possible outcomes for 3 coin-flips

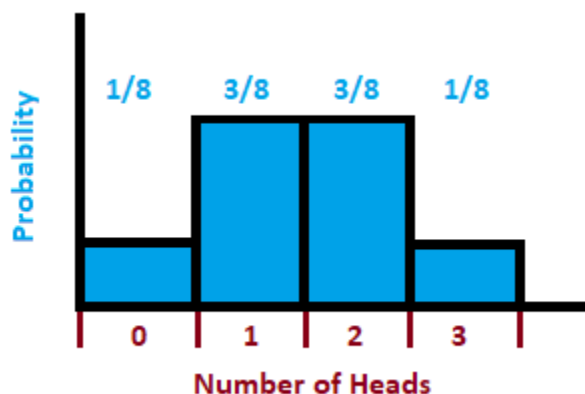


Figure 22: Graphing the probabilities of the outcomes of three coin flips

As we perform more coin flips, the probabilities form a bell shape that approximates the normal distribution, shown in Figures 23 and 24.

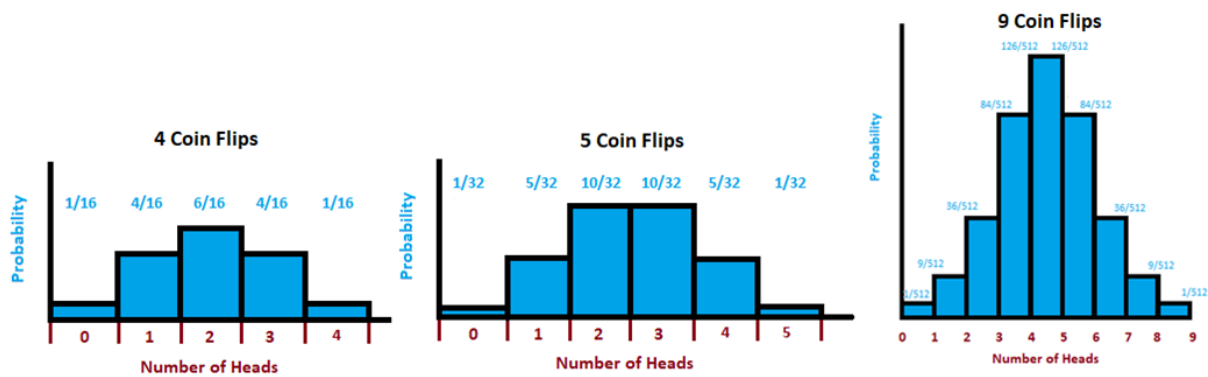


Figure 23: Graphing the probabilities for 4, 5, and then 9 coin flips.

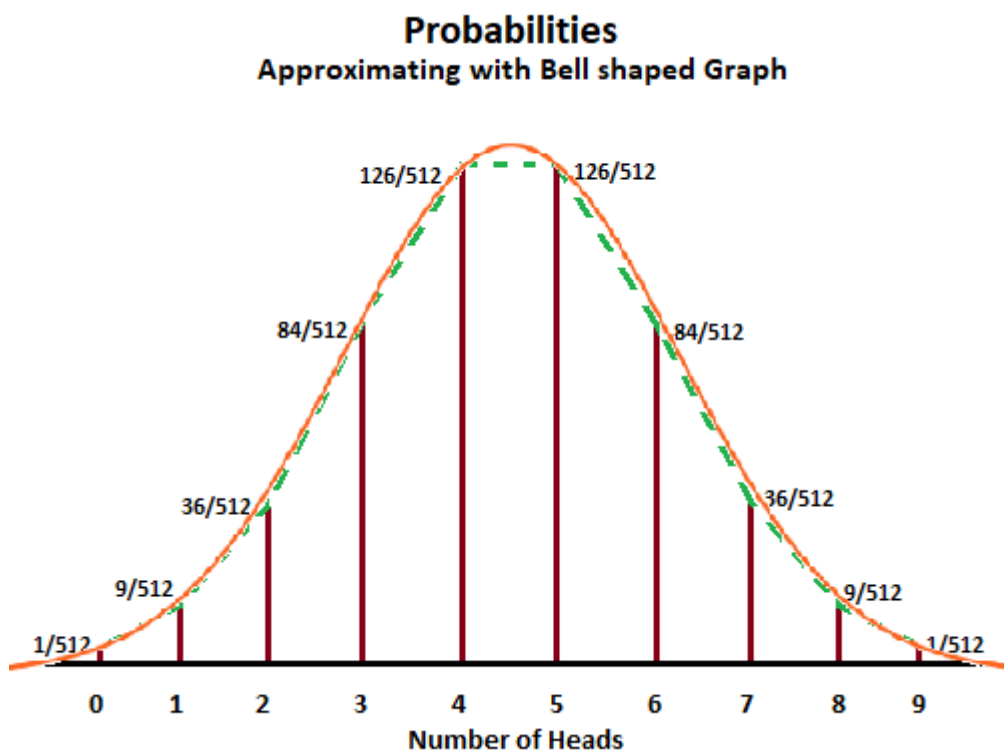


Figure 24: Probabilities of getting a certain number of heads in 9 coin flips. Note the bell shape.

Note that the probabilities for flipping a coin 9 times are the same as the probabilities for the 9th row of a Galton Board. Just as in the Galton Board example, flipping more coins (similar to adding more rows of pegs) will create a smoother bell curve, with the highest probabilities near the center and the lowest probabilities on the more extreme ends as illustrated in Figure 25.

C. Driving Example

A car is driven in a city with a grid similar to the one shown in Figure 27, and the driver is only allowed to travel southbound or eastbound.

Question: How many different paths can the driver take from point A to point B?

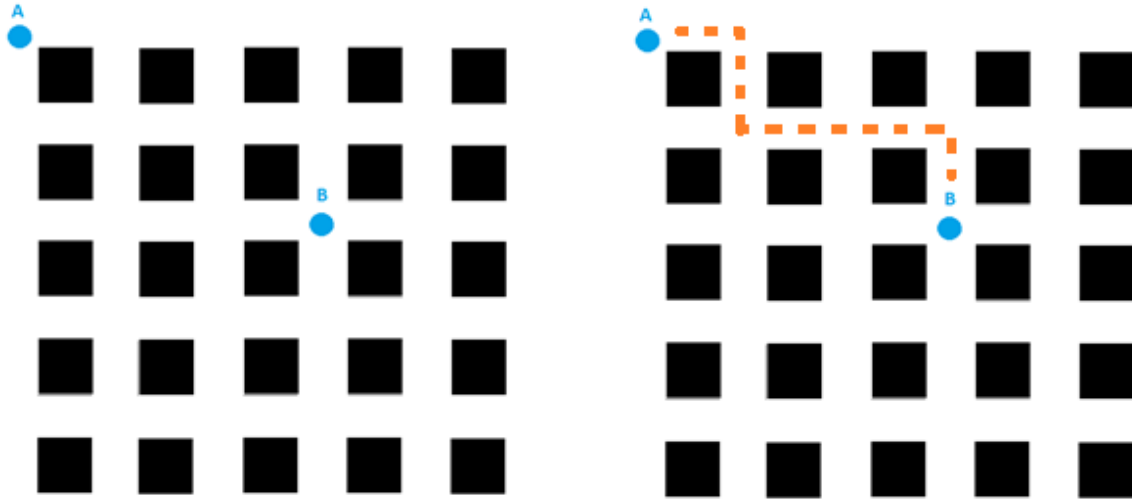


Figure 27: Two points in a city grid: A is the starting point; B is the final destination (left), and an example of one possible path from A to B (right)

The purple dots shown in Figure 28 mark destinations that only have one possible path from point A. Destination X has 2 possible paths from point A, down and to the right, or to the right and then down. Destination Y has 6 possible paths from point A, and Z has 20 possible different paths from point A.

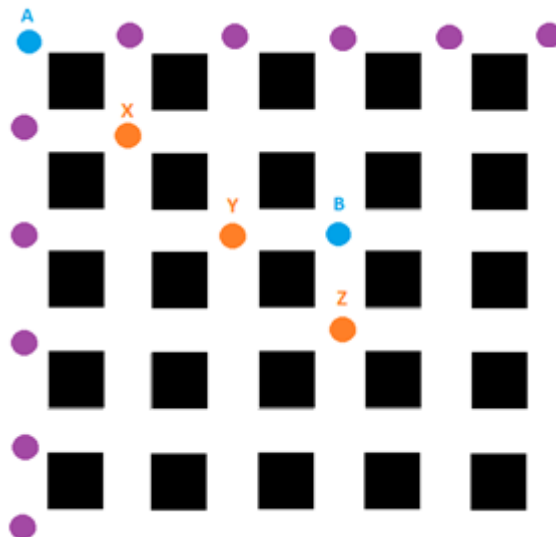


Figure 28: A city grid with key points marked

Figure 29 shows the number of possible paths from point A to any other destination on the grid. So, the answer to the original question is that there are 10 different possible driving paths from point A to point B.

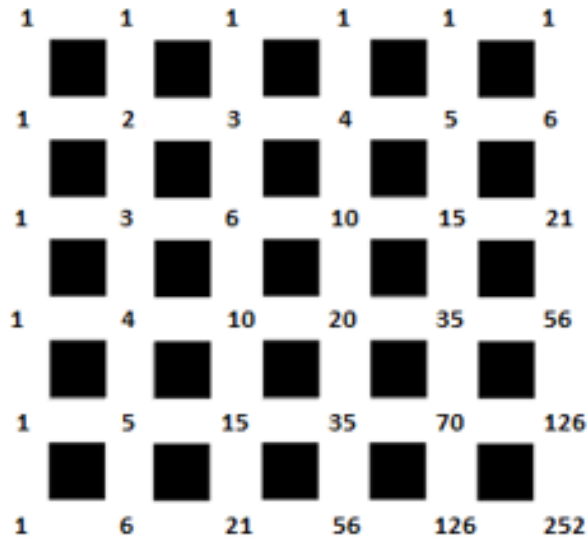


Figure 29: City grid with number of possible paths shown

If we turn the grid diagonally and look at each row, as in Figure 30, a familiar pattern emerges.

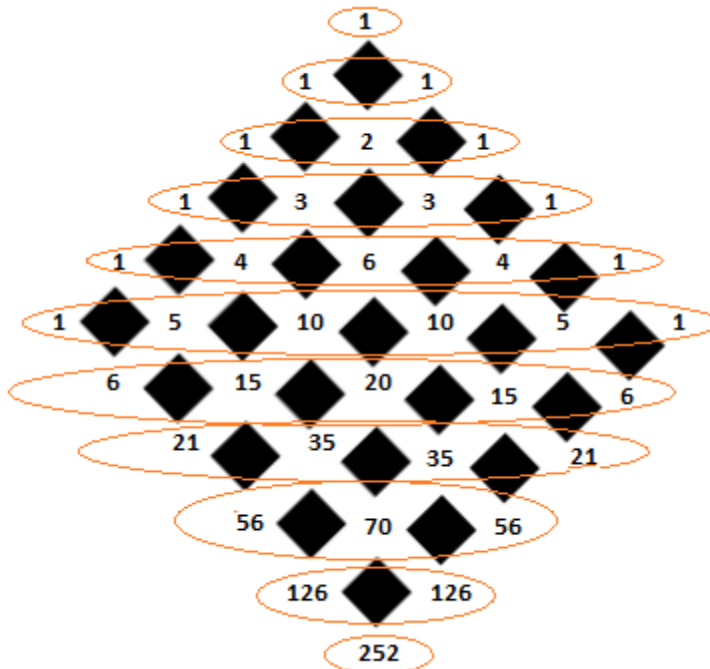


Figure 30: The grid, turned diagonally, shows a familiar pattern.

This pattern is Pascal’s Triangle. Note that this pattern, shown below in Figure 31, is the same as in the case of the coin flip experiment and Galton Board example.

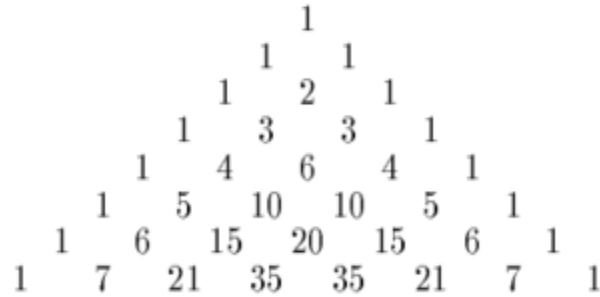


Figure 31: Pascal’s Triangle

III. Introducing Key Concepts

A. What Does a Bell Curve Mean?

The shape of a bell curve tells us about the probability of given outcomes. The higher sections have a higher chance of occurrence, so the “middle” outcomes are more likely (recall the near center columns in a Galton Board) and the extreme left and right (outlier) outcomes are the least likely. When viewing Figure 32, the blue highlighted area is related to the probability of that specific region, just like a specific outcome on a Galton Board or a particular number of “heads” outcomes in the coin flip experiment.

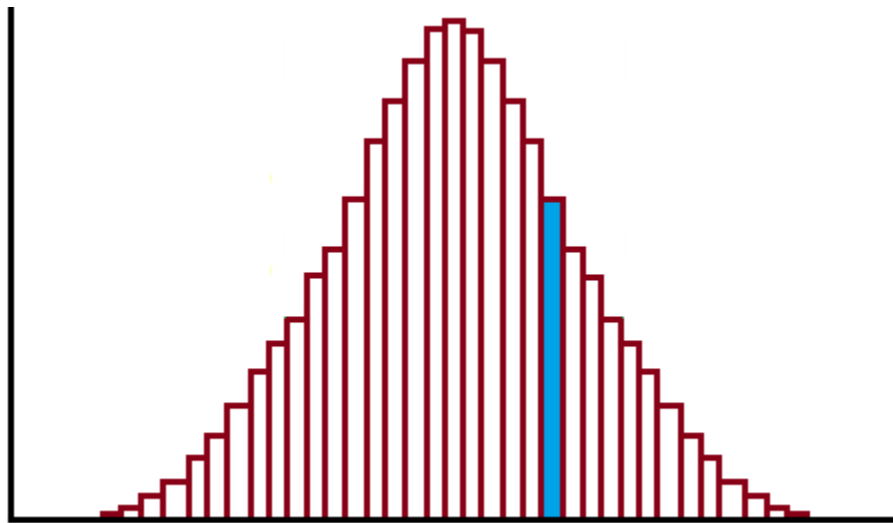


Figure 32: The blue area represents the probability of a specific outcome.

More generally, we can find the probability of range of desired outcomes by integrating (taking the area under the bell curve between two points), as shown in Figure 33. While there are

a few different types of bell curves, the normal distribution is the most common and the examples shown in this paper approximate the normal distribution.

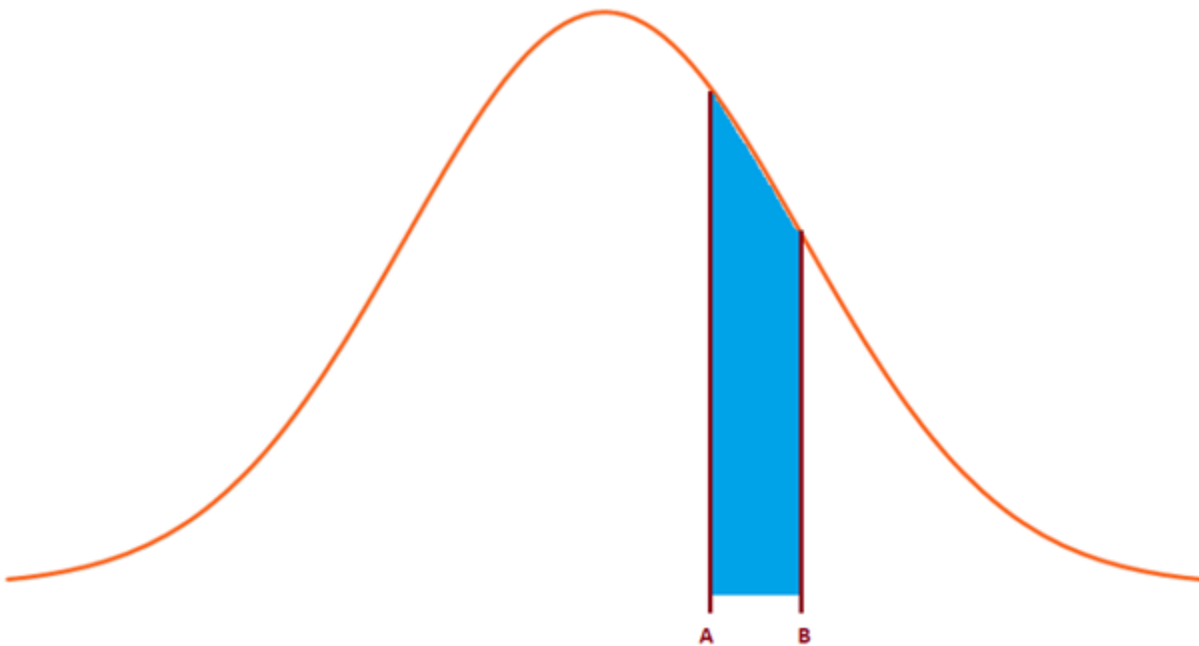


Figure 33: The shaded area represents the probability of getting outcomes in the range between A and B.

B. More on the Normal Distribution

The normal distribution is a commonly used probability distribution, but it should be noted that not all bell curves are normal distributions. A graph of a normal distribution is a symmetrical bell-shaped curve that has its mean at the highest point, as shown in Figure 34, but also has a precise formula. The Figure also shows the standard deviation as related to the normal distribution.

The pink section in Figure 34 is one standard deviation to the right and to the left of the mean. Approximately 68.2% of all data samples lie in this region. The pink and blue regions together represent two standard deviations in either direction from the mean, and represents approximately 95.4% of the data samples. Three standard deviations in either direction represents about 99.7% of all data samples.

The integral of the normal distribution curve, or the area under the curve, always equals to 1. This simply means that the probability of all possible events added together is 100%.

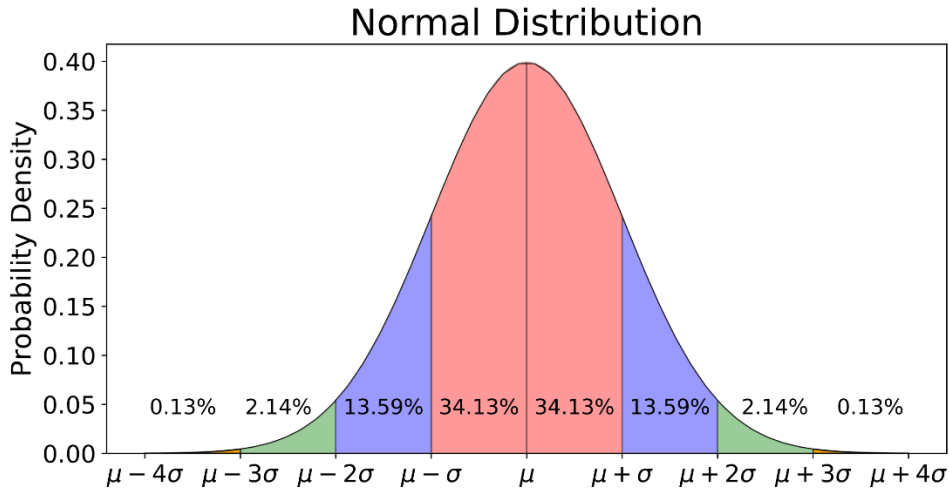


Figure 34: Normal distribution and related probabilities

Source: <https://towardsdatascience.com/understanding-the-68-95-99-7-rule-for-a-normal-distribution-b7b7cbf760c2>

IV. Real-Life Examples of the Normal Distribution

A. Height of People

Suppose a class has just started and one student is running late. What would students guess the height of this random student to be, assuming you know they're male? Most students would likely guess a height somewhere between 5'5" (65 inches) and 6'2" (74 inches), but it's doubtful that anyone would guess the student to be 8' tall. Why is that?

In the United States, male height appears to form a normal distribution [10] with the average being 5 feet, 9.3 inches (69.3 inches) and with a standard deviation of 2.94 inches. This means that about 68.2% of adult men in the US have a height between 5' & 6.36" and 6' & 0.24". Very few people are either very tall or very short. Assuming we can plot it in a continuous manner, it is illustrated with the normal distribution in Figure 35.

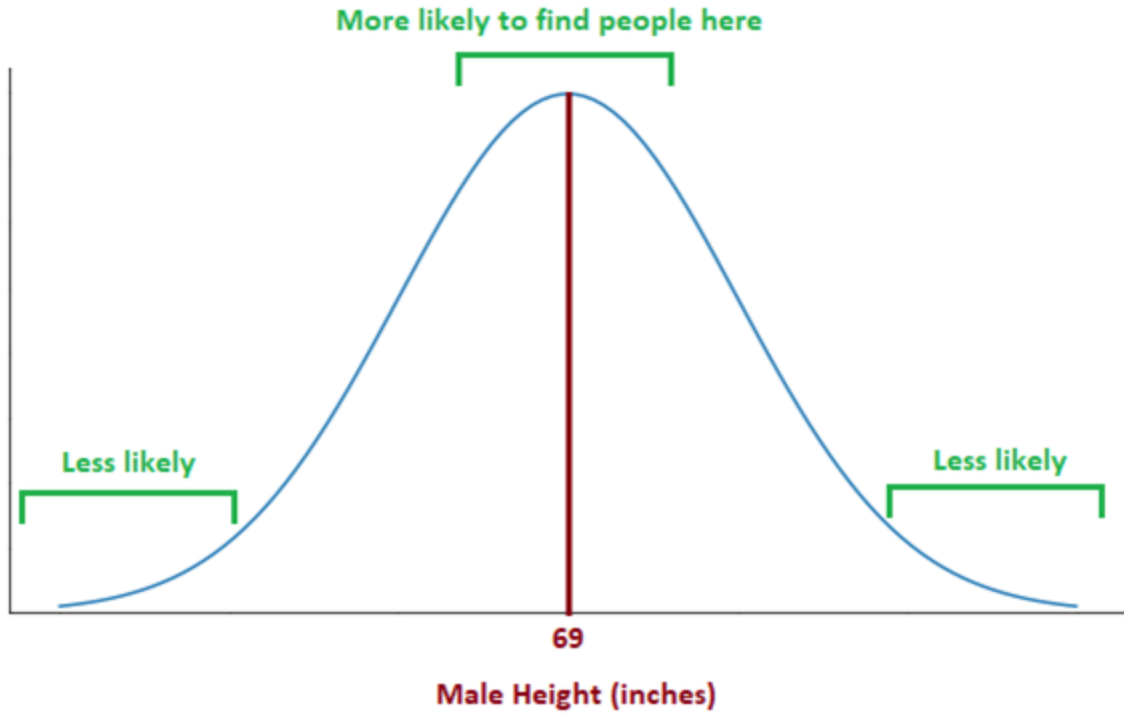


Figure 35: Chart showing male height. More people are near an “average” height while less people are very tall or very short.

The total area under the curve equals 1, as shown in Figure 36, which means that any person “has a height” is a sure event.

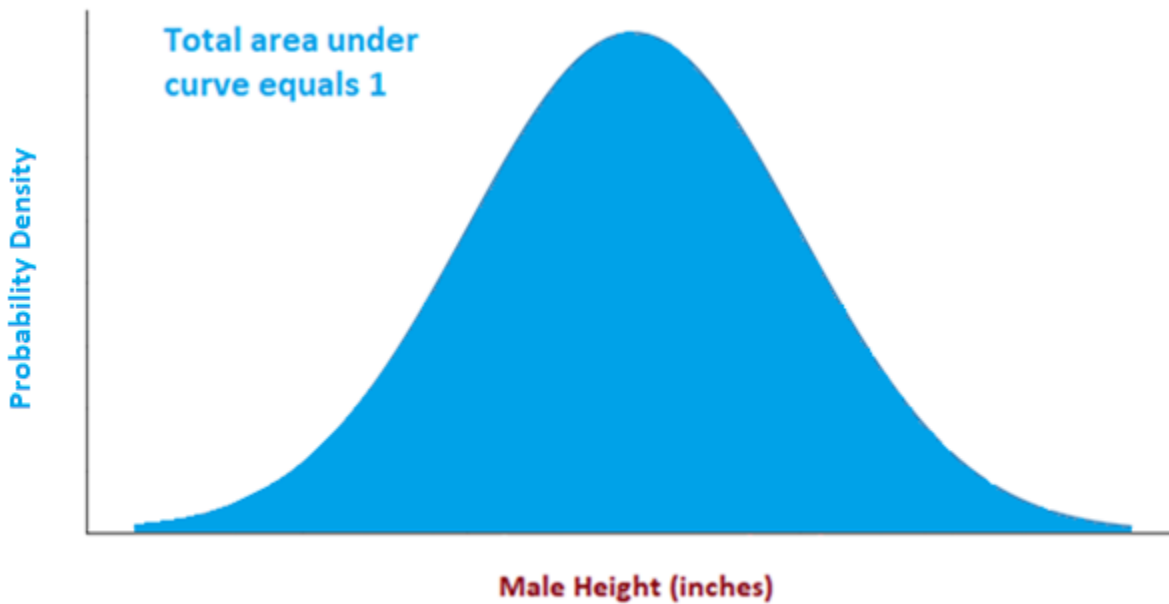


Figure 36: The total area under the curve equals 1

If we highlight and focus on one portion of the curve, like the yellow one shown in Figure 37, the area in that region represents the probability of someone having a height between the two values (for this figure, 74 and 80 inches).

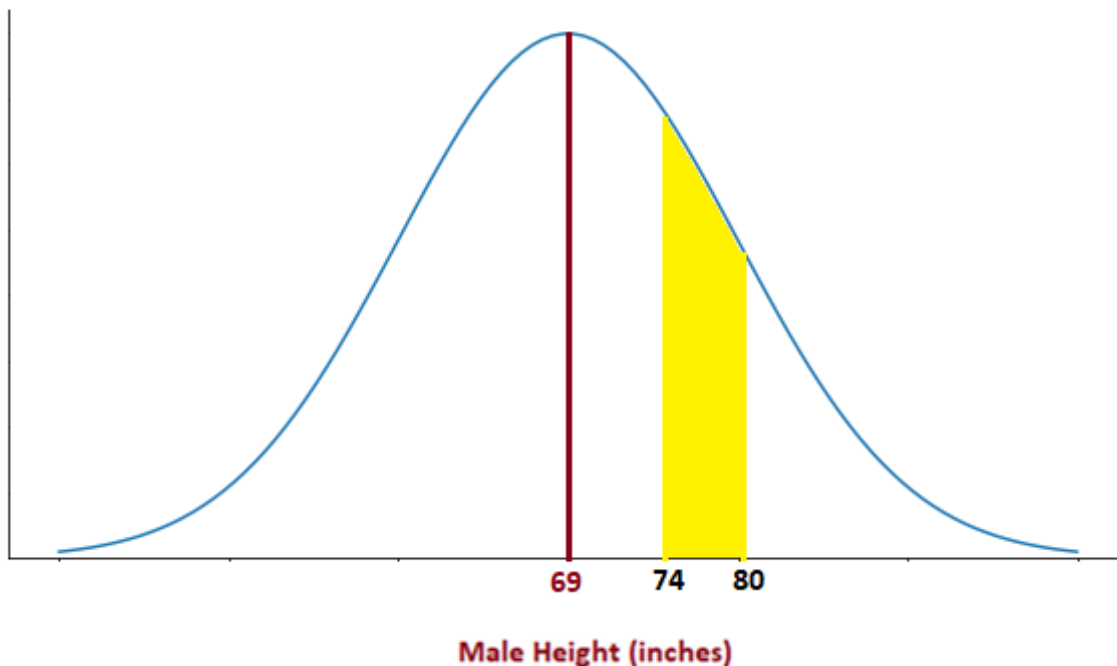


Figure 37: The area of the yellow region is the probability of someone having a height between the values of 74 and 80 inches.

B. Intelligence and IQ

Intelligence Quotient, commonly known as IQ, is a way to measure certain human intelligence. It is measured through the use of a few different standardized tests. Current IQ tests are developed so that the mean IQ is 100 and the standard deviation is 15 IQ points. This means about 68.2% of all people have an IQ between 85 and 115 (that is, one standard deviation from the mean in either direction), and about 95% of people have an IQ between 70 and 130 (two standard deviations in either direction from the mean). It is important to note that IQ is developed and adjusted so that the data will fit a normal distribution.

V. Puzzle – Multiple Choice Exam

Suppose you get to class one morning and find out that you have to take an exam you had forgotten about. You didn't study at all for the exam, but there are 20 questions and each question has two options: true and false. If you select answers randomly, how many questions do you expect to get right?

Since each question has two options, the chance of randomly selecting the right answer on any given question is 50%, or 0.5. Since we only care about how many questions are correct, and not which particular questions are correct, the order of the results doesn't matter (i.e. for our purposes, getting only the first ten questions correct is the same as only getting the last ten questions correct).

The average number of questions guessed correctly is the probability for each question multiplied by the number of questions, or 0.5×20 . Thus, if you simply guess for each question you should expect to get 10 questions right. If we graph the probabilities for all 20 questions, it forms an approximation to normal distribution, since a true/false exam is really a binomial problem like the coin flip, and the sum of the probabilities is 1.

VI. In-Class Experiment – Rolling a Die

A single die roll is similar in concept to a coin-flip, except that here there are 6 possible outcomes instead of two. Let us consider rolling a die 10 times.

When you roll a die once, you have an equal ($1/6$) probability for each number one through six. If you roll 10 dice, the sum of the ten outcomes will be a number between 10 (assuming all ones) and 60 (assuming all sixes). The most likely outcomes are closer to values around 35, since the “middle” outcomes have more possibilities.

VII. Assessment

The methods and examples shown in this paper were presented and assessed in a Probability and Statistics class taught by the first author of this paper at Florida Atlantic University. A questionnaire was distributed to students at the end of the lecture in order to gauge how receptive students were to the intuitive and engaging learning techniques presented in this paper. The responses are based on a class presentation that included examples from this paper. 21 students filled out the questionnaire. A summary of the questionnaire can be found below, and the raw results can be found in Appendix A.

Figures A.1 and A.2 show the self-identified demographic breakdown of participating students. There is a fair amount of diversity, which is representative of the diversity at Florida Atlantic University and its engineering programs. Most students were 23 years old or younger, but there were a handful of older students.

Questions regarding student opinions about importance of the lecture or learning in general had five options: very important, important, moderately important, slightly important, and not important. Questions regarding opinions about effectiveness of various presentations or learning methods had the options: strongly agree, agree, neutral, disagree, or strongly disagree.

When asked about the “importance of being introduced to Probability/Statistics through visual examples,” one-third of the students responded with “very important” and an additional

one-third responded with “important”. Only four students marked “slightly important,” and one student marked “not important.”

Students were generally unfavorable on their opinions of the importance of traditional and PowerPoint presentations. When asked about the “importance of learning using traditional presentations,” only 3 students marked important and none marked very important. Nine students responded moderately important, while nine more responded with either slightly important or not important. Answers to the question regarding the importance of learning Probability/Statistics through PowerPoint presentations were similar.

Students subjectively found that visual presentations helped them understand the concept of the normal distribution. This is apparent by the responses to the question: “How effective was the presentation on Probability/Statistics today?” and “What could have helped you to understand the concept of the normal distribution better?” most students wrote that they found the presentation effective, specifically the visual portions of the presentation.

These results, when taken as a whole, show that many students feel that they learn best through visual and engaging approaches. They generally are neutral to the idea of traditional and PowerPoint presentations, and prefer to be taught rather than teach themselves.

VIII. Conclusion

This paper attempts to help students develop intuition towards the concept of the normal distribution by using a visual, intuitive, and engaging approach based on real-life examples. Multiple examples have been shown with the hope that instructors may find them useful and incorporate them in their teaching. It is not expected that instructors will use this work in its entirety, but instead that they will pick and choose the examples that they believe their students will benefit the most from. We hope that by using these examples more students will experience the “Aha” moment and these concepts will become clear to them.

Results of the assessment, shown in Appendix A, show very positive. Students felt that these examples helped them to learn and understand the normal distribution in a clear and intuitive way.

IX. Acknowledgements

The authors thank Professor Aalo for allowing us to teach and assess this concept in his class. Also, the Center for Writing at Florida Atlantic University provided great feedback.

X. References

- [1] P. Peebles, "Probability, Random Variables, and Random Signal Principles," McGraw-Hill Science/Engineering/Math, 2000.
- [2] S. Takahashi, "The Manga Guide to Statistics," No Starch Press, 2008.
- [3] B. Orlin, "Math with Bad Drawings," Black Dog and Leventhal Publishers, 2018.
- [4] D. Huff, "How to Lie with Statistics," W. W. Norton & Company, 1993.
- [5] S. Campbell, "Flaws and Fallacies in Statistical Thinking," Dover Publications, 2004.
- [6] A. Gelman and D. Nolan, "Teaching Statistics: A Bag of Tricks," 2002.
- [7] D. Rumsey, "Statistics for Dummies," For Dummies, 2016.
- [8] C. Wheelan, "Naked Statistics," W. W. Norton & Company, 2014.
- [9] J. Phillips, "How to Think About Statistics," Holt Paperbacks, 1999.
- [10] J. Cook, "Distribution of Adult Heights," *johndcook.com*, 2008.

Appendix A – Assessment Results

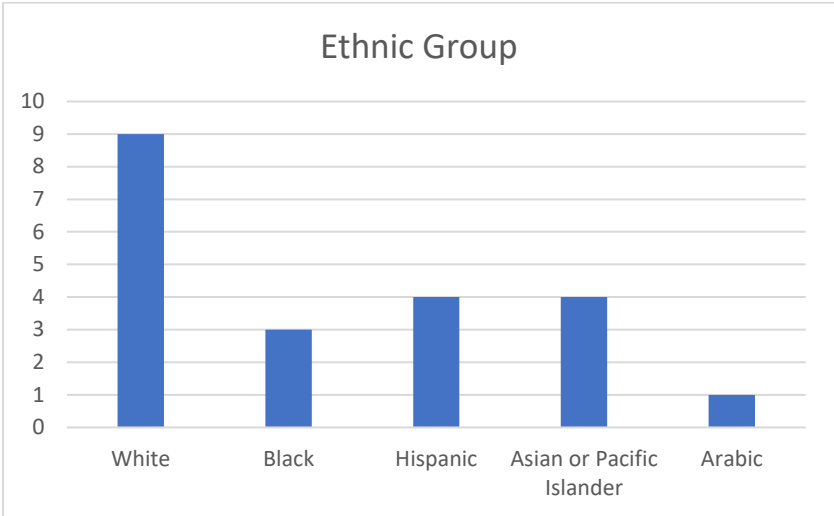


Figure A.1: Self-identified ethnic distribution

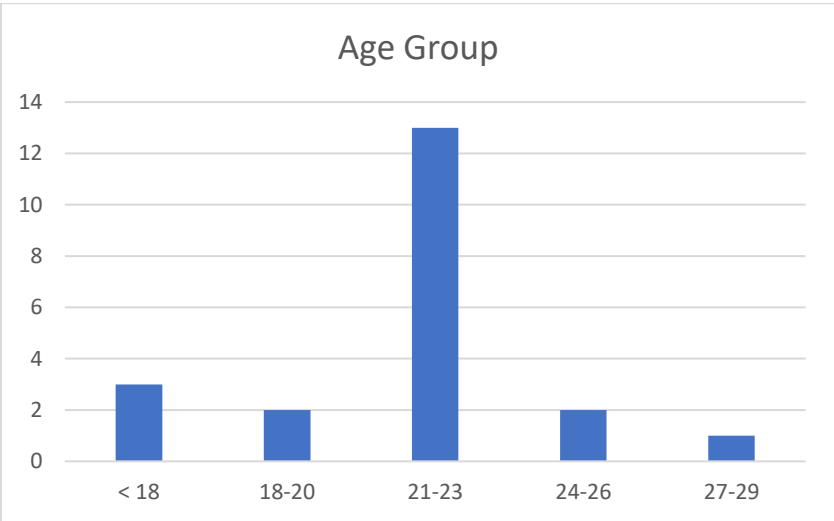


Figure A.2: Self-identified age distribution

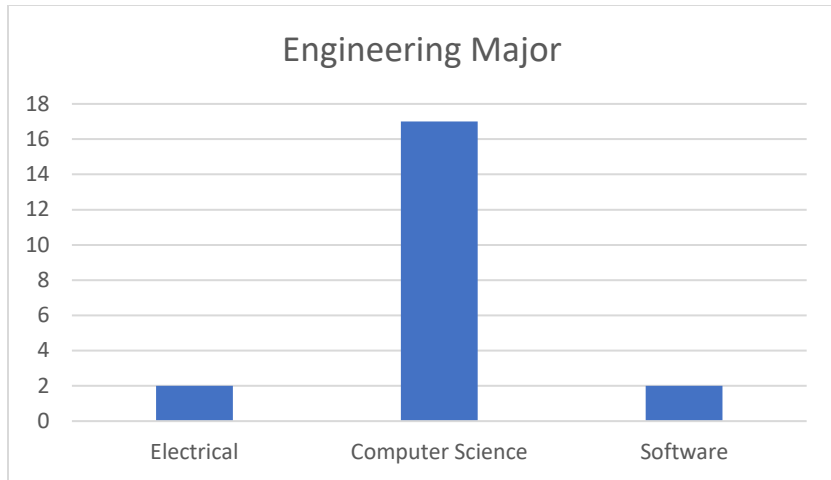


Figure A.3: Engineering Major

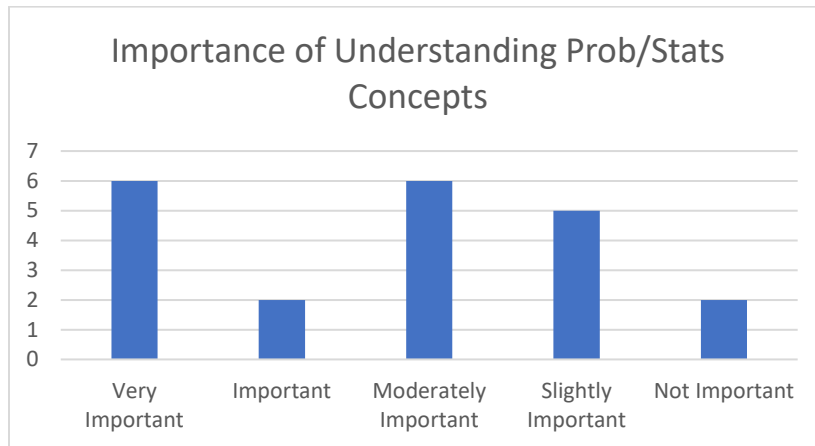


Figure A.4: Student feedback on importance of understanding Probability/Statistics concepts

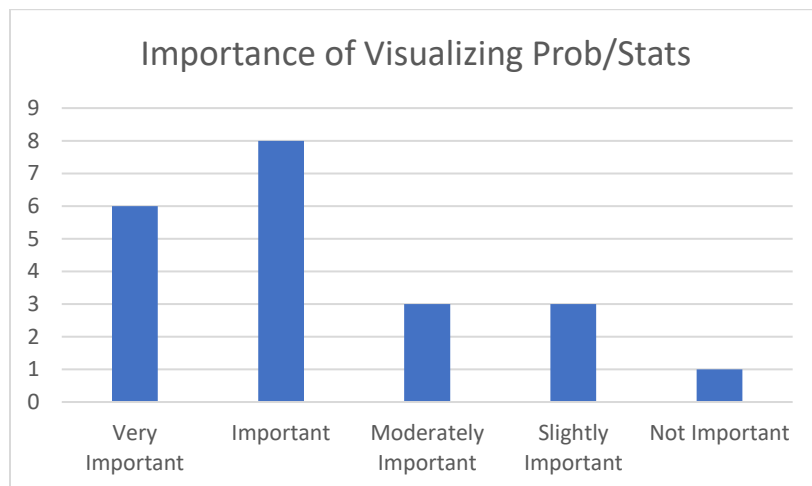


Figure A.5: Student feedback on importance of visualizing Probability/Statistics

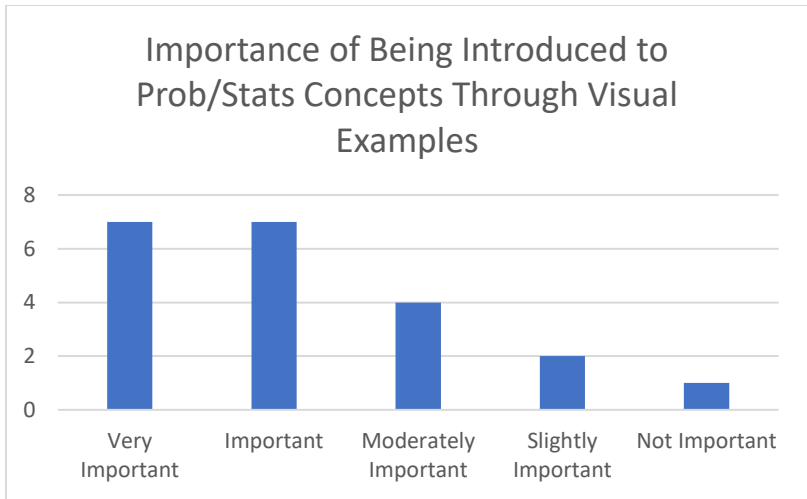


Figure A.6: Student feedback on importance of being introduced to Probability/Statistics concepts through visual examples

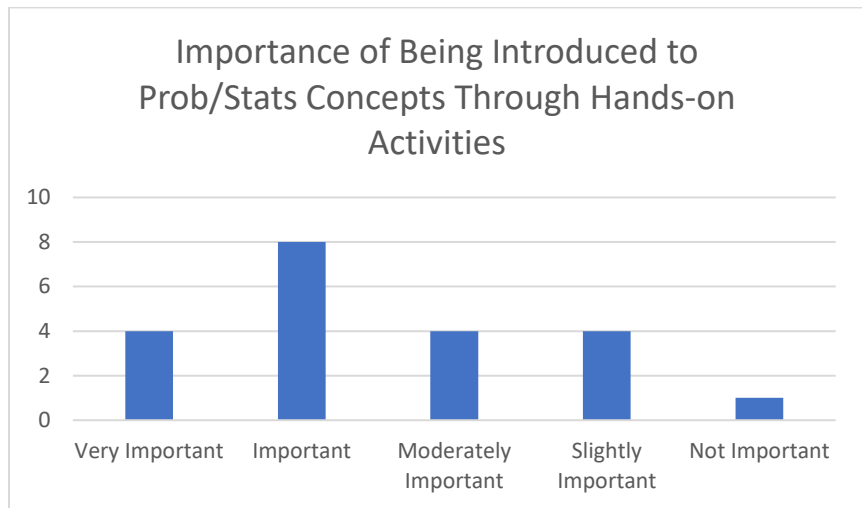


Figure A.7: Student feedback on importance of hands-on activities in Probability/Statistics

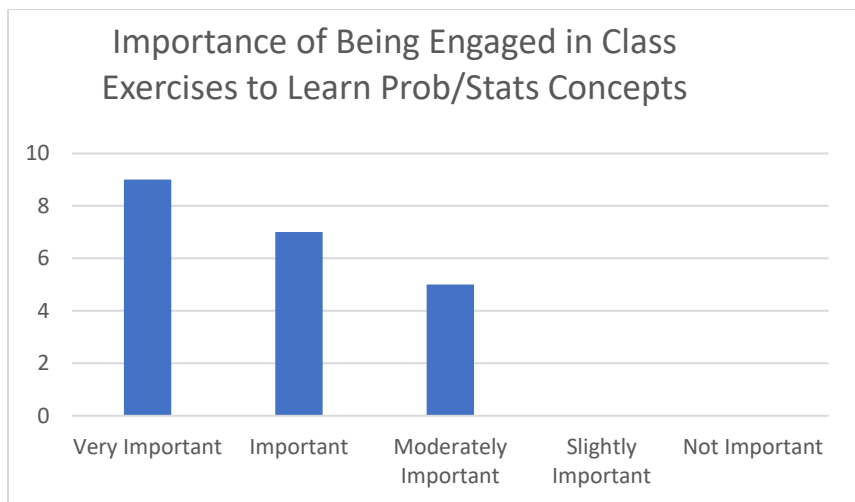


Figure A.8: Student feedback on importance of class engagement in Probability/Statistics

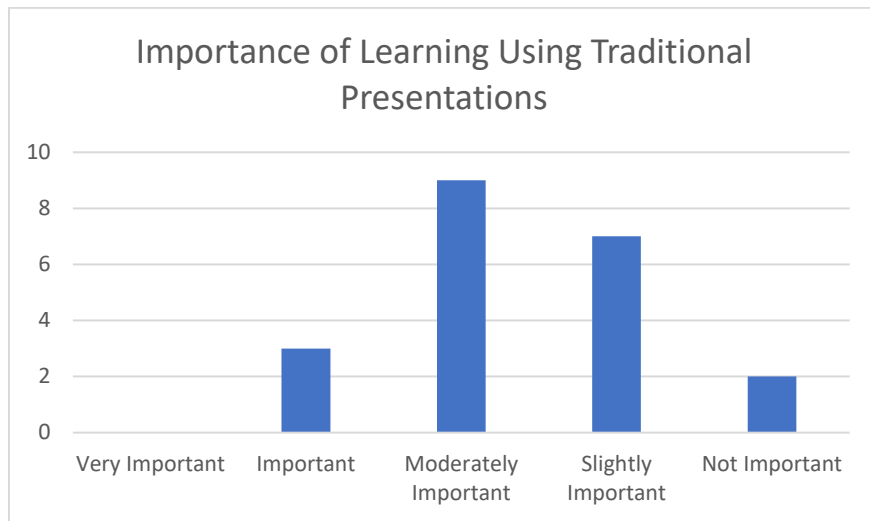


Figure A.9: Student feedback on importance of traditional presentations for learning

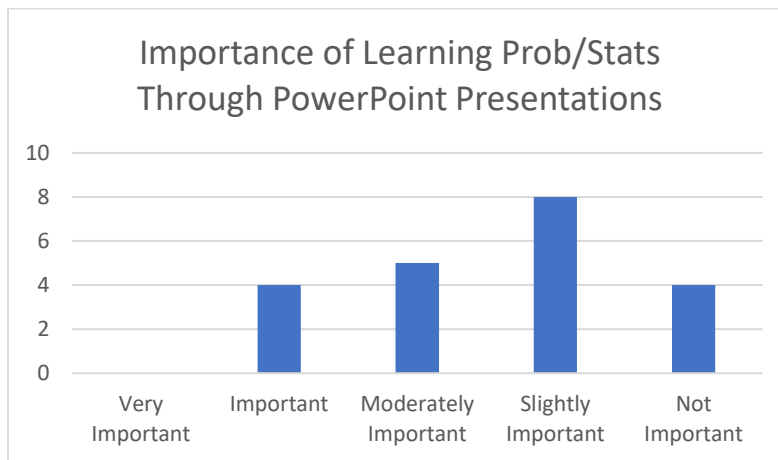


Figure A.10: Student feedback on importance of PowerPoint presentations for learning Probability/Statistics

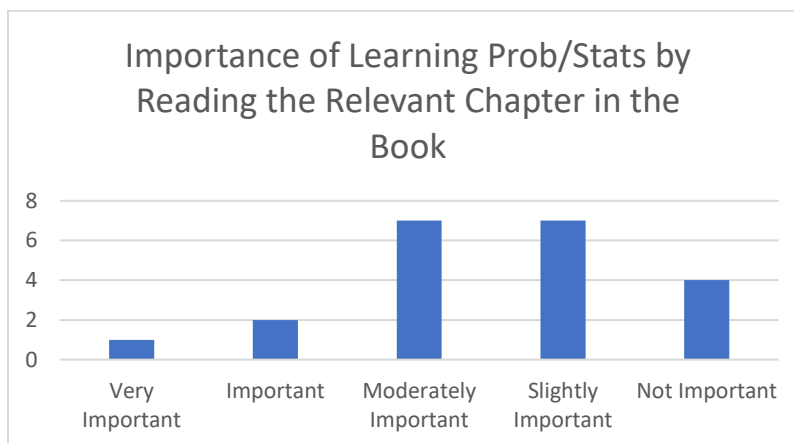


Figure A.11: Student feedback on importance of learning Probability/Statistics through textbook reading

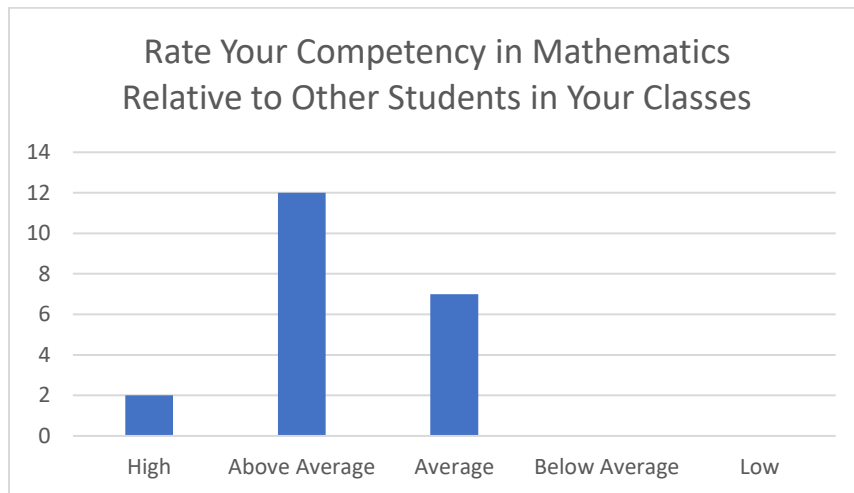


Figure A.12: Student self-assessment of mathematics competency

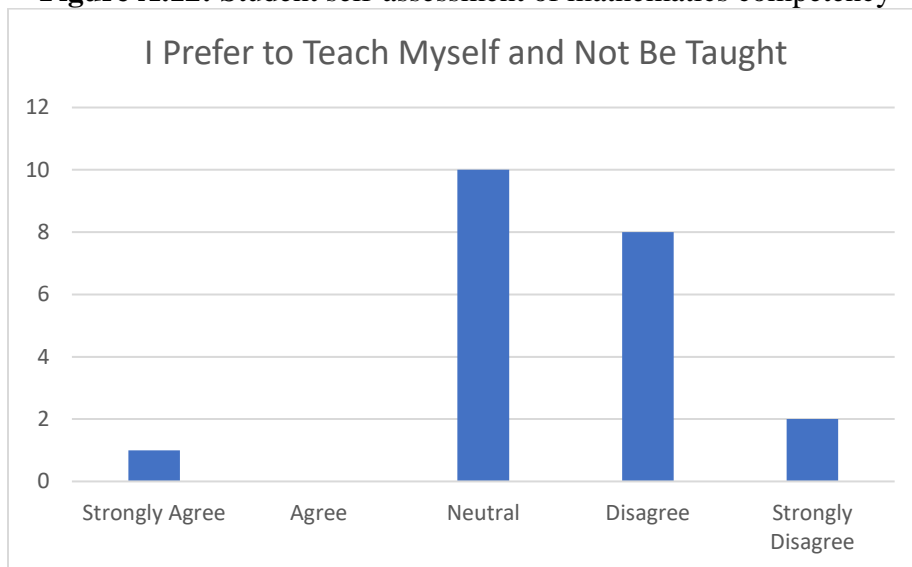


Figure A.13: Student opinions on self-teaching

Appendix B – A Few Notes on Pascal’s Triangle

1. Notable Features

Pascal’s Triangle in the simplest sense, is a particular number pattern. Pascal’s Triangle has several notable features. The first is the way in which diagonal rows increase. The first diagonal row is all ones. The second is counting starting from one. The third row is triangular numbers (1, 3, 6, 10, etc.). Note that since the triangle is symmetrical, this is true for starting from either the left or the right.

Another notable feature is that this number pattern can be used to find the probability for problems like the location of a marble on a Galton Board, coin flip probabilities, or the driving example (more specifically, it can be used to find probabilities for binomial problems). To do this, select the number at the location where you want to know the probability and divide by the sum of the numbers in that row. The sum of any row is also 2 raised to the power of the row number minus one. For example, the probability for a marble to arrive at the middle location in row 5 of a Galton Board is $6/(2^4)$ or $3/8$.

As another example, to find the probability of getting a number of heads for 2 coin flips, look at the third row. The options are 0, 1, or 2 heads results. Since the numbers in the third row are 1, 2, and 1, this means the probability of getting 0 heads is $1/(1+2+1) = 1/4 = 0.25$. The probability of getting 1 heads result is $2/4 = 0.5$, and the probability of getting 2 heads results is also 0.25.

2. Combinations

Figure B.1 shows a different way to write the numbers of Pascal’s Triangle.

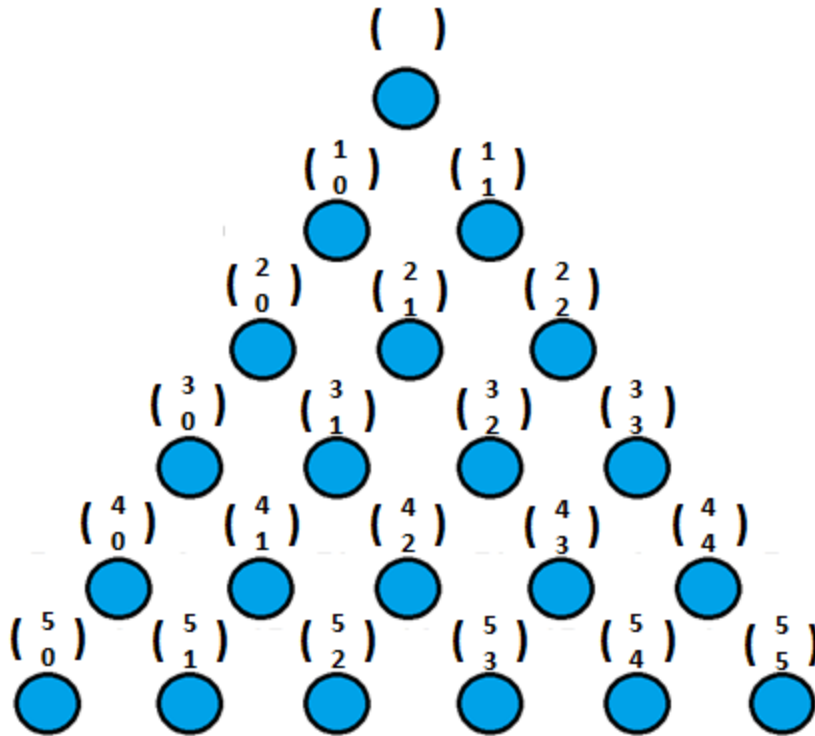


Figure B.1: Pascal's Triangle representing combinations instead of numbers

The symbols shown in Figure B.1 show “combinations,” and represent the chances of success, using the following equation:

$$\binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

This equation solves for the probability of “x” successes in “n” trials where “p” is the probability of success. When we see $\binom{3}{1}$ as a chance of success, it means 1 success in 3

trials, or $\frac{3!}{1!(2!)}$ which is 6/2 or 3, so there are 3 opportunities for that outcome. When multiplied by the chance of success (50% success / 50% failure for the Galton board or coin flips), we get $3 * (0.5)^3$ which is 0.375. So, for example, the probability of getting 1 “heads” result in 3 coin flips is 0.375.

Since the sum of each row is equal to 2^{n-1} , shown in Figure B.2, we should also note that:

$$\binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{n-1} + \binom{n}{n} = 2^n$$

$$\begin{array}{ccccccc}
 & & & & \binom{0}{0} & & = & 2^0 \\
 & & & & \binom{1}{0} & \binom{1}{1} & = & 2^1 \\
 & & & & \binom{2}{0} & \binom{2}{1} & \binom{2}{2} & = & 2^2 \\
 & & & & \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & = & 2^3 \\
 & & & & \binom{4}{0} & \binom{4}{1} & \binom{4}{2} & \binom{4}{3} & \binom{4}{4} & = & 2^4 \\
 \binom{5}{0} & \binom{5}{1} & \binom{5}{2} & \binom{5}{3} & \binom{5}{4} & \binom{5}{5} & = & 2^5
 \end{array}$$

Figure B.2: Each row is equal to 2^{n-1}

It should also be noted that since, as previously discussed, each number is the sum of the two numbers above it, we can show the following, shown in Figure B.3:

$$\begin{array}{c}
 \binom{3}{1} \quad \binom{3}{2} \\
 \swarrow + \searrow \\
 = \binom{4}{2}
 \end{array}$$

Therefore, more generally

$$\begin{array}{c}
 \binom{n}{r} \quad \binom{n}{r+1} \\
 \swarrow + \searrow \\
 = \binom{n+1}{r+1}
 \end{array}$$

Thus,

$$\binom{n}{r} + \binom{n}{r+1} = \binom{n+1}{r+1}$$

Figure B.3: A general case for combinations in Pascal's Triangle

3. Polynomials

A pattern that derives from the previous figure, shown in Figure B.4, can help us find the probability for any point on Pascal's Triangle. For each level (n), starting with 0, the probability for any point is $x * (0.5)^n$ where x is the number from Pascal's Triangle. Note that the sum of all probabilities in any given level is equal to one. That is to say, some outcome at each level is guaranteed.

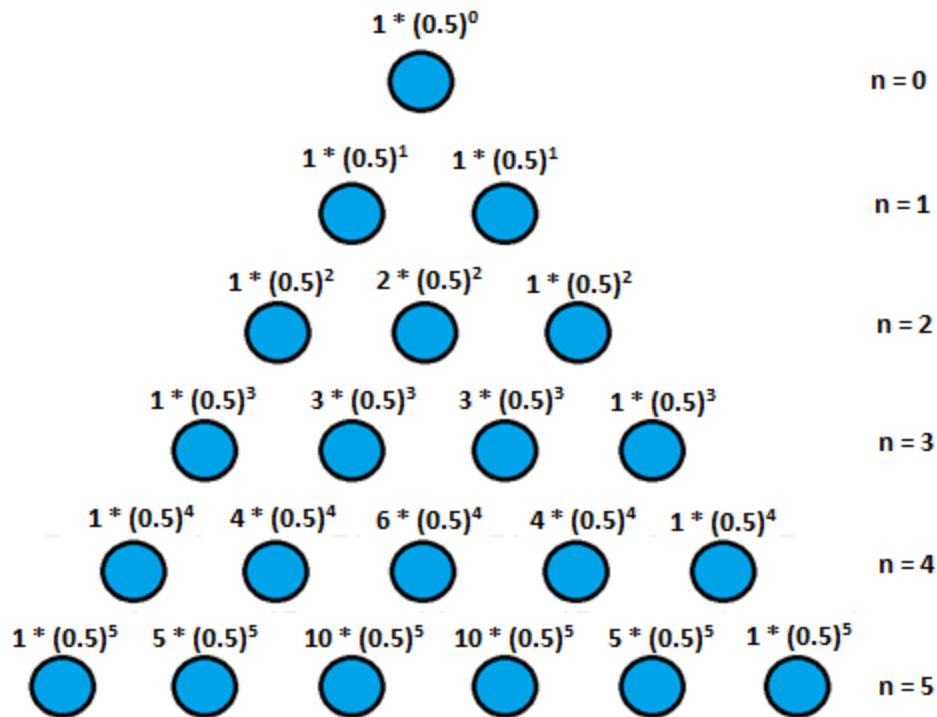


Figure B.4: The probability for any point on Pascal's Triangle is given by $x * (0.5)^n$.

Consider the expression $(a+b)^n$, where n counts upward from zero. Figure B.5 shows these expanded expressions up to $n = 4$, and for the general case n . When we consider the coefficients of each term, we see that it once again forms Pascal's Triangle.

$$(a + b)^0 = 1$$

$$(a + b)^1 = a + b$$

$$(a + b)^2 = a^2 + 2ab + b^2$$

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a + b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

$$(a + b)^n = a^n + na^{n-1}b + \dots + na^1b^{n-1} + b^n$$

Figure B.5: Expansion of the expression $(a+b)^n$. Note that the coefficients are "binomial" and represent Pascal's Triangle.

So why do all coefficients add up to 2^{n-1} ?

$$a + b = \binom{1}{0} a^1 b^0 + \binom{1}{1} a^0 b^1$$

Has two coefficients, $2 = 2^1$

$$(a + b)^2 = a^2 + ab + ba + b^2$$

Has 4 coefficients that we "condensed" to 3. $4 = 2^2$

$$(a + b)^3 = a^3 + a^2b + a^2b + b^2a + a^2b + b^2a + b^2a + b^3$$

Has 8 coefficients that we "condensed" to 4. $8 = 2^3$

Similarly, $(a+b)^n$ has 2^n coefficients that can be condensed to $(n+1)$ coefficients.