



Number 87

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS

The Statistics Ceacher Network

ASA/NCTM Joint Committee on Curriculum in Statistics and Probability

The Editor's Corner

The Statistics Teacher Network



Angela Walmsley

Love STEW lessons? Have a great idea you'd like to share? Join us as a reviewer or author and contribute to our collection of highquality statistics lessons. We're looking for lessons at all levels: K–5, 6–8, and 9–12. Contact: *steweditor*@ *amstat.org*

www.amstat.org/ education/stew I hope everyone is having a wonderful winter quarter, semester, or trimester—but I know I'm personally looking forward to spring and some warmer weather! I'm really excited to share with you our next edition of the Statistics Teacher Network—Issue 87.

The first article is by one of our associate editors, Doug Rush from Saint Louis University. We thought we would try a "throwback Thursday" item in this edition, similar to what you might see on social media sites such as Facebook. Doug has chosen the "throwback" item—the famous Monty Hall problem from "Let's Make a Deal." It reminds us of the common mistake made during this fun game on probability.

The second article is by Octavious Talbot, Alex Ocampo, Sam Tracy, Kelly Mosesso, and Marcello Pagano in the department of biostatistics at the Harvard T.H. Chan School of Public Health and 10th-grade student Aisha Rahman from the Boston Latin School. This article describes a unique program, StatStart, developed by the group and targeted at minority highschool students. The program is based in biostatistics and provides a summer opportunity to introduce statistical concepts through the lens of biostatistics and the programming language R. Various problem sets are offered—including an adaptation of the Monty Hall problem that is the focus of our first article!

Our last article is by Sue Haller and Melissa Hanszek-Brill from St. Cloud State University. This article describes a hands-on approach to teaching the mean and standard deviation concepts (and why the algorithms work the way they do). Using square tiles, the authors describe multiple samples of teaching standard deviation alongside the traditional algorithm. The activities presented here will allow students to hold a mental image for the process of calculating standard deviation and connect the algorithm to a concrete model.

I encourage you to read all these great articles! All have information that can be used directly in the classroom at various levels.

We encourage our readers to write for the *Statistics Teacher Network*! We love to publish a variety of articles about statistics throughout the Pre-K–16 range. Our editors are also happy to assist if you have an idea for an article. Please send any articles or ideas you have for consideration to *Angela*. *Walmsley@cuw.edu*.

Regards,

Angela Walmsley, Editor Concordia University Wisconsin

Articles

Let's Make a Deal: The Monty Hall Problem	.2
StatStart 2015	.4
A Conceptual Approach to Teaching Standard Deviation	.9

ASA PR Campaign for Students, Teachers,	
Counselors Continues Expanding Reach	14
Featured STEW Lesson Plan	16
Announcements	37

Spring 2016

Throw Back: Let's Make a Deal: The Monty Hall Problem

By Douglas K. Rush, Saint Louis University

The game show "Let's Make a Deal" originally aired on NBC television in 1963 with Monty Hall as its creator, producer, and longtime host. The show has had enduring popularity and continues to run on CBS with current host Wayne Brady who, when not hosting LMD, stars as Lola—the red-sequined, thigh-high, stiletto-heeled-boot-wearing drag queen in the Broadway hit musical "Kinky Boots."

In addition to its popularity with television audiences, "Let's Make a Deal" has given rise to a popular probability puzzle commonly referred to as the Monty Hall Problem. On the show, Monty Hall would invite a contestant to join him on stage. In front of them stood three doors. Monty would tell the lucky contestant that behind one of the three doors was a brand new car (or similar great prize). However, behind the other two doors were goats (or as I recall from one episode, chickens).

Monty would then ask the contestant to pick a door. For the purpose of this article, we will assume the contestant picks door number one initially. After the contestant picks door number one, Monty opens door number three and reveals a goat. Monty does not then famously ask, "Do you want to keep door number one, or would you like to switch to door number two?" However, the problem has had so much attention that most people think he asked about switching the doors. (*www. youtube.com/watch?v=c1BSkquWkDo*). What if he had asked? Should the contestant keep door number one or change to door number two?

This was the point at which the probability paradox arose. The solution was so counterintuitive that it did not seem correct. The correct choice based on conditional probability is that the contestants should always switch doors from thier initial pick. Let's examine the probability behind this solution.

The initial probability of picking the correct door is where the probability of an event = the number of ways the event can occur, divided by the number of possible outcomes. Figure 1 depicts the initial probability of selecting the door that hides the car as 1/3. There are three doors that can be selected, but the car is only behind one of the doors. Conversely, the initial probability of selecting a door that hides a goat is 2/3, as two of the three doors hide goats.

Many would believe changing doors does not improve the odds of winning the car. They would argue that the probability of correctly picking the door hiding the car at that point is 1/2,



Figure 1 Initial Door Selection Probability



Figure 2 Probability of Winning After Changing Doors

since there are now two unopened doors, only one of which hides the car. The fallacy of this logic is that we have new information related to the initial pick. Originally, we only had a 1/3 probability of selecting the door hiding the car, but we had a 2/3 probability of selecting a door hiding a goat.

Monty has now revealed one of the doors hiding the goat. If we switch doors, we have a 2/3 probability of winning the car. See Figure 2. Of course, this does not mean we win the car if we switch doors. There is still a 1/3 probability that the car was hidden behind door number one. What it does mean is that there is a 2/3 probability that we win the car by switching doors.

I know there are skeptics out there. As I tell my statistics students, I am not a real scientist; I am a social scientist. Nevertheless, I decided to conduct a series of trials to test the Monty Hall Problem. Not having doors, goats, or new cars at my immediate disposal, I decided to let three common playing cards represent the three doors. Two of the cards were deuces (2) and represented the unwanted goats. One of the cards was an ace and represented the new car.

I decided I would conduct 100 trials. My graduate assistant played the role of Monty. Monty shuffled the three cards and laid them, face down, on top of my desk. I determined in advance that I would alternate my pick of cards in a 1, 2, 3 sequence. In other words, I would pick the first card in the first trial, I would pick the second card in the second trial, and I would pick the third card in the third trial. In the fourth trial, I would pick the first card and would continue with the 1, 2, 3 sequence until I had run 100 trials.

I also determined in advance that, after my initial card pick, I would always change cards after Monty revealed that one of the unpicked cards was a deuce, an incorrect choice. I recorded whether the change resulted in winning the car (picking the ace) or selecting the goat (picking the deuce). After 100 trials, the change of cards resulted in 65 wins (ace/car) versus 35 loses (deuce/goat). Incredibly, almost exactly the 2/3 probability we expected!

This can serve as a fun project to demonstrate conditional probability to your class. Create several two-student teams. One student on each team will be Monty and the other student on the team will be the contestant. Using playing cards, have the Montys shuffle three cards and lay them face down and instruct the contestants to select a card. Each Monty should then inspect the cards and turn over one of the unpicked cards that hides the goat. Each contestant should then change cards and have each Monty reveal whether the change resulted in winning the car or the goat. Have each team run 10 trials and record their results. After the 10 trials, add all the teams' trials results to determine the overall probability of winning the car after changing from the first pick.

The Monty Hall Problem has been debated for many years and is the subject of many discussions about probability. An excellent historical discussion of the various statistical bases for the Monty Hall Problem can be found at *https://en.wikipedia.org/wiki/Monty_Hall_problem#Bayes.27_theorem*. A good short video of the problem can also be found on the Khan Academy website. ■

References

https://en.wikipedia.org/wiki/Monty_Hall_problem #Bayes.27_theorem (pictures in the article are from the Wikipedia site)

www.khanacademy.org/math/precalculus/prob_comb/ dependent_events_precalc/v/monty-hall-problem

www.youtube.com/watch?v=c1BSkquWkDo

StatStart 2015

By Octavious Talbot, Alex Ocampo, Sam Tracy, Kelly Mosesso, and Marcello Pagano of the Harvard T.H. Chan School of Public Health Department of Biostatistics and Aisha Rahman of the Boston Latin School

The lack of interest and diversity in students pursuing science, technology, engineering, and mathematics (STEM) subjects has been a persistent issue at all levels of education, and, consequently, in the workforce. Concerned that a lack of interest and exposure to statistics at the school level may contribute to this, a group from the department of biostatistics at the Harvard T.H. Chan School of Public Health founded StatStart, a program targeted to high-school students and designed to increase the STEM applicant pool at the college level and thus diversify the field.

StatStart was intended as a month-long intensive computing and biostatistics course geared toward high-school students from under-represented minority and low-income backgrounds in the local Boston area. The program met throughout the month of July 2015 for four hours a day, four days a week at Harvard.

The philosophy behind this summer program was to attract young high-school students at an age at which they are most likely open to exploring, appreciating, and learning different intellectual skills. In particular, we wanted to expose them to something they were unlikely to come across elsewhere in their schooling: computer programming and biostatistics.

As languages tend to be more easily learned when young, it should be possible to teach a computer language to young students—the younger the student, the more natural a language becomes. To make the language more interesting and attractive, we decided to use biostatistics as a vehicle. This was a subject they would most likely be unfamiliar with, but we thought they would find it exciting and of tremendous social relevance. With a more clearly defined solution to our goals, the next step was to build such a program.

To launch StatStart, we composed admissions criteria, publicized the program to local high schools in the Boston area, and invited a small number (10) of students. At first, the students felt a little overwhelmed due to their limited prior knowledge of either biostatistics or programming. However, we helped them build a foundation in these areas by having them explore various programming techniques in R (a free statistical programming language) through class work, group work, and homework. The students were progressively more eager to learn as their knowledge of and interest in the subject grew. Quickly realizing that they were surrounded by kindred spirits, they didn't even mind the homework!

While collaboration was heavily encouraged, each participant was responsible for turning in his/her individual solutions. The students voluntarily contacted one another outside of class through social media and video chats, where they relayed their notes and explanations of the concepts. This consequently enhanced their understanding of the topics at hand, as well as lending insight into related problems. From this, stemmed a tight-knit camaraderie. Ayub Ahmed, the only male StatStart 2015 student, recalls, "As days went by, every single person from the StatStart became my friend, actually best friends, and ... we became connected [even] outside class."

A typical class started with an in-class demonstration/lecture by one of the graduate students, followed by a back-andforth discussion with the students regarding their progress and the homework from the previous day. The feedback from these sessions was used to create future homework tailored to reinforce skills and topics students struggled with. We also strived to incorporate time for students to work on problems in groups and individually during class, where we could provide immediate support as needed.

The curriculum developed by the StatStart team provided a rigorous introduction to both programming and statistical concepts. Students mastered important foundational programming skills—including loops, if/then statements, and functions—while also learning how to use R for statistical purposes such as performing exploratory data analysis, linear regression, and hypothesis testing.

In addition to class work, we had students of the graduate program, including members of under-represented minorities, talk to the StatStarters during lunch about how each of them had come to be in Harvard's biostatistics department. What was perhaps surprising to the high-schoolers was that few of the lunch-time speakers had knowledge of biostatistics prior to college. These speakers emphasized how valuable exposure to the field during their pre-collegiate years would have been. The graduate students also spoke of their research, which served to show students other areas of statistics. Carla, a second-year doctoral student, presented elements of population genetics and genome-wide association analysis in addition to her own motivations for statistical study and interests in genetics. We thought showing the high-schoolers role models who were just a few years older than they were would make STEM careers seem more achievable and interesting. In addition, we also thought the graduate students would benefit from explaining their research to nonstatisticians!

A primary goal of ours in beginning StatStart was to increase the pool of under-represented minority students pursuing STEMrelated fields of study in college, so our program would have been incomplete without addressing the college application process. Therefore, we had the students attend one of Harvard's admissions information sessions, where they learned about the admissions process and campus life from current students and an admissions counselor. Students also had the opportunity to have lunch with current undergraduate students who were in the department for the annual Summer Program in Biostatistics and Computational Biology (*www.hsph.harvard.edu/ biostatistics/diversity/summer-program*). This allowed our high-schoolers the chance to discuss application-related matters with current college students.

As StatStart came to a close, each graduate student instructor was paired with groups of two high-school participants to work on a comprehensive project. The projects were designed to be completed by the end of the course and culminated in presentations on their methods and results to a room full of friends, family, and statisticians. The quality of the presentations and depth of understanding the students showed on complex issues such as the Law of Large Numbers, the Central Limit Theorem, and the concept of herd immunity made us all proud. These topics are attached at the bottom of this article.

One group used the weak law of large numbers (WLLN), ran a large number of iterations, and gave insight into questions such as "What is the probability that two individuals in a random group of 24 have the same birthday?" and "What is the probability of three random individuals on an elevator going to consecutive floors when there are 10 floors?" In probability theory, WLLN describes the results of performing the same experiment or game a large number of times. A description of students' work on these tasks follows.

Law of Large Numbers (LLN)

In probability theory, the law of large numbers (LLN) is a theorem that describes the result of performing the same experiment a large number of times. According to the law, the sample average of the results obtained from a large number of trials should be close to the "true" distribution mean and will tend to become closer as more trials are performed. The LLN is important because it "guarantees" stable long-term results for the averages of some random events. For example, while a casino may lose money in a single spin of the roulette wheel, its earnings will tend toward a predictable percentage over a large number of spins. Any winning streak by a player will eventually be overcome by the parameters of the game. It is important to remember that the LLN only applies (as the name indicates) when a large number of observations is considered. There is no principle that a small number of observations will coincide with the expected value or that a streak of one value will immediately be "balanced" by the others (see the gambler's fallacy).

This theorem can be demonstrated easily via computer simulation. Using 10,000 iterations, the students coded the following scenarios and gave their corresponding probabilities of occurring:

a. Given a random sample of 24 people, what is the probability that two share a birthday? Three? Assume

that each day of the year is equally likely to be someone's birthday and ignore leap years.

- b. Three people get into an empty elevator on the first floor of a building that has 10 floors. Each presses the button for their desired floor (unless one of the others has already pressed that button). Assume they are equally likely to want to go to floors 2 through 10. What is the probability that the buttons for three consecutive floors are pressed? Now suppose there are four people in the elevator and one will definitely go to the fourth floor. Recalculate the probability for three consecutive floors.
- c. The Monty Hall Problem: Suppose you're on a game show and you're given the choice of three doors. Behind one door is a car; behind the others, goats. You pick a door, say No. 1. The probability that you have chosen the door with the car is 1/3. Now the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" What is the probability of winning the car if you choose door No. 2?

According to the law, when the outcome is binary (success or failure), the average of the results obtained from a large number of trials should be close to the actual probability of success. In computing, this corresponds to defining success as "1" and failure as "0." We store the results after each experiment and conduct the identical experiment a large number of times making sure the random structure is still preserved as to provide a variety of outcomes. Because success is defined as 1 and failure 0, we will have the count of the number of successes if we use basic addition and add all the 1s and 0s from the experiments. Therefore, dividing that sum by the total number of experiments gives us the proportion of experiments that were successful. If we perform this experiment 10,000 times, it adopts the interpretation of the probability of success.

The birthday problem follows:

Given a random sample of 24 people, what is the probability that two of them share the same birthday? Assume each day of the year is equally likely to be a random individual's birthday and ignore leap years.

To tackle the birthday problem, the students coded a pseudoexperiment on the computer in which 24 numbers were randomly selected with replacement from a list of numbers 1:365. For each iteration of the experiment, if the same number appeared twice, the computer interpreted the results as two individuals being born on the same day and stored a 1 in an empty vector. If not, a 0 was stored. After running 10,000 iterations and collecting the results of each iteration, the students summed up each of the 1s and 0s and divided by the total number of iterations, providing an accurate estimate of the probability. The students explored this question in depth by running the experiment with 10, 100, 1,000, and 10,000 iterations several times each and gained further insight into the prowess of the weak law of large numbers. The group concluded there was an estimated 54% chance that two random individuals in a group of 24 had the same birthday.

The elevator question was slightly more involved as the students had to invoke algebraic skills learned in previous classes to give insight to the question:

Three random people get into an empty elevator on the first floor of a building that has 10 floors. Each presses the button for their desired floor (unless one of the others has already pressed that button). Assume they are equally likely to want to go to floors 2 through 10. What is the probability that the buttons for three consecutive floors are pressed?

Again, the students coded a pseudo-experiment on the computer and ran it for 10,000 iterations. During each iteration, the students randomly sampled three numbers with replacement from 2:10 to represent the three individuals going to random floors. Because it was sampled with replacement, it allowed multiple people to go to the same floor. Once the three numbers were sampled, the students instructed the computer to take the difference between the maximum number and the median number, say X, and the median number and the minimum number, say Y. Using algebra, the students proved the numbers are in consecutive order if and only if X = Y = 1. Therefore, if this was the case, a "1" was stored, else a "0" was stored after each iteration. After every iteration was run, the students added all the 1s and 0s and divided by 10,000 to conclude that the estimated probability three patrons go to consecutive floors is 5.6%.

The Monty Hall Problem, which was the most involved to code, follows:

Suppose you're on a game show and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1. The probability that you have chosen the door with the car is ¹/₃. Now the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" What is the probability of winning the car if you choose door No. 2?

This is a controversial question due to typical intuition leading individuals to believe there is a 50/50 chance; however, due to the host knowing which door has the car and intentionally opening the goat door, we gain useful information that actually increases the probability that you win the car by switching to roughly 66%. The students verified this result using simulations at 10,000 iterations. During an iteration, a vector containing two "os" and one "1" is initially constructed—in this vector, the "1" represents the car and the "os" represent the goats. Next, a random position is chosen and that number is turned into a "5," regardless of whether it was initially a "1" or "0." This represents the door that the contestant initially guesses. Now, the "os" that remain in the vector are counted. If there is one "0" remaining, then the contestant initially chose a goat door and the host has no choice but to remove the final goat door. However, if there are two "os" left, then that would imply the contestant initially chose the car door and the host has flexibility in what door he removes.

In the computer, the positions of the two goat doors are stored, a random number between 0 and 1 is generated, and the door in the first position is removed if that number is less than 0.5, else the door in the second position is removed. A goat door is removed in the computer by turning a "0" into a "2."

Finally, the contestant has to decide whether to switch; however, we are interested in how often the contestant wins by switching. In this effort, we check to see if "1" remains in the vector after identifying the door guessed and the door removed. If it remains, that would imply the contestant did not guess the car door, nor did the host remove it. Therefore, if the contestant switches, he/ she would win the car. If this is the case, a 1 was stored during the iteration, else a 0. The students verified the estimated probability of winning the car via switching to be 66%.

In conclusion, the students exhibited a deep understanding of the material that reached far beyond our expectations. We cannot help thinking the marriage between a computer language and statistical theory for high-schoolers was beneficial. The motivating wishes of StatStart were surpassed, and we firmly believe it will have a critical broader impact on the greater Boston community.

StatStart directly addresses the education gap by strengthening the college application of the participants and advances the knowledge of the STEM field through diverse recruitment. It also provides like-minded high-school students from various backgrounds a place to learn, assist one another, and achieve while gaining experience and exposure to STEM. It was a successful first year, and we hope to make it an essential part of the Harvard T.H. Chan School of Public Health's identity as the program evolves and expands.

Although our program was offered primarily for minority students, it can be readily implemented for any high school and for any group of students. Having this alternative also allows better experiences for students studying statistics in a more formal environment because it provides an applied perspective grounded in an experiential setting to promote increased interest in the subject—students from our program continue to contact us for help searching for more opportunities in biostatistics.

We recommend structuring future offerings in an analogous fashion as was used in StatStart; namely, tailor the content to best suit the students in the program, of course, and then find real applications that are motivating and challenging. Such content is easy to generate when the teachers are well versed in the application area. If such teachers are not available, then an effort should be made to collaborate with a practicing statistician who can work with the teacher(s) to present the field in an understandable manner to school-aged children.

Appendix: Other project topics

Central Limit Theorem

The normal distribution is used to help measure the accuracy of many statistics, including the sample mean, using an important result called the Central Limit Theorem. This theorem is one of the most famous in statistics and gives you the ability to measure how much the means of various samples will vary without having to take any other sample means to compare it with. By taking this variability into account, you can use your data to answer questions about a population, such as "What's the mean household income for the whole U.S.?" or "This report said 75% of all gift cards go unused. Is that likely?" (These two particular analyses are made possible by applications of the Central Limit Theorem called confidence intervals and hypothesis tests, respectively.)

This theorem can be demonstrated using computer simulation.

- a. Given that you continuously try an experiment of your choice in which you have a success or failure every time you run it, choose a probability of success greater than 0, but less than 1. You repeat this experiment until you have five successes. Count the number of times you need to run the experiment. Repeat this exercise and draw a histogram of the number of times you need to get your five successes. This histogram should be bell shaped. Study how many times you need to repeat the exercise before you are willing to call the histogram bell shaped.
- b. Given a normal 52-hand deck of cards, you draw 10 cards from the deck. We are interested in the expected number of aces drawn in these randomly selected 10 cards. Let's say you repeat this drawing of 10 cards 100 times and calculate the mean number of aces. Show that the mean of the number of aces drawn approximates a normal distribution despite coming from something called the hypergeometric distribution.

c. You are being tested for psychic powers. Suppose you do not have psychic powers. A standard deck of cards is shuffled, and the cards are dealt face down one by one. Just after each card is dealt, you name any card (as your prediction). Let X be the number of cards you predict correctly. Assuming you get no feedback after each prediction until all cards have been flipped, experiment with several guessing strategies, pick your favorite, and show that the average number of correct cards is normally distributed given your selected strategy.

Herd Immunity

Herd immunity (also called herd effect, community immunity, population immunity, or social immunity) is a form of indirect protection from disease that occurs when a large percentage of a population has become immune to an infection, thereby providing a measure of protection for individuals who are not immune. In a population in which a large number of individuals are immune, chains of infection are likely to be disrupted, which stops or slows the spread of disease. The greater the proportion of individuals in a community who are immune, the smaller the probability that those who are not immune will come into contact with an infectious individual.

Individual immunity can be gained through recovering from a natural infection or through artificial means such as vaccination. Some individuals cannot become immune due to medical reasons. In this group, herd immunity is an important method of protection. Once a certain threshold has been reached, herd immunity will gradually eliminate a disease from a population. This elimination, if achieved worldwide, may result in the permanent reduction in the number of infections to zero, called eradication. This method was used for the eradication of smallpox in 1977 and for the regional elimination of other diseases.

The effects of herd immunity can be seen through computer simulation.

a. Consider a community of 10,000 people in which 95% are vaccinated against a certain contagious disease. One person with the disease enters the population and comes into contact with 50 random people in the community. If the community member is vaccinated, there is a 0% chance of transmission. If the community member is not, there is a 60% chance that the person will contract the disease. Assuming all individuals who have the disease or contract the disease come into contact with 50 people randomly and spread the disease with the same probabilities listed, simulate this scenario until the rate of new infections is o. After the simulation is complete, calculate the prevalence of the disease. Explore different parameters for this study (e.g., probability of infection, number of people in the population, vaccination rate). Be sure to keep track of the parameters used and the results.

Infants Data Set

Birth weight is the first weight of your baby, taken just after he or she is born. A low birth weight is less than 5.5 pounds. A high birth weight is more than 8.8 pounds. A low birth weight baby can be born too small, too early (premature), or both. This can happen for many reasons. They include health problems in the mother, genetic factors, problems with the placenta, and substance abuse by the mother. Some low birth weight babies may be more at risk for certain health problems. Some may become sick in the first days of life or develop infections. Others may suffer from longer-term problems such as delayed motor and social development or learning disabilities. High birth weight babies are often big because the parents are big or the mother has diabetes during pregnancy. These babies may be at a higher risk of birth injuries and problems with blood sugar.

A low birth weight can be caused either by a preterm birth (low gestational age at birth), the infant being small for gestational age (slow prenatal growth rate), or a combination of both. The goal of this project will be to justify this fact using the infants data set of all African-American and Caucasian births in North Carolina. To do so, please answer the following scientific questions. Justify your answer using statistical hypothesis testing:

- a. Are African-American babies smaller than Caucasian babies?
- b. Is gestational weeks associated with birth weight?
- c. Controlling for all information in the data set, which of the factors collected significantly affect birth weight?
- d. Controlling for all information in the data set, which of the factors collected significantly affect death of the baby before 365 days?

While this project uses a data set used extensively in class, the methods used to analyze it will extend far past what has been done in class so far. Students will use statistical methods that are learned during the first year of the biostatistics PhD program.

World Bank Data Set

The World Bank is a vital source of financial and technical assistance to developing countries around the world. It is not a bank in the ordinary sense, but a unique partnership to reduce poverty and support development.

The World Bank Group has set two goals for the world to achieve by **2030**:

- End extreme poverty by decreasing the percentage of people living on less than \$1.25 a day to no more than 3%
- Promote shared prosperity by fostering the income growth of the bottom 40% for every country

For this project, the World Bank data set has information about life expectancy, fertility rate, population, and GDP per capita for almost every country in the world for every year since 1960. There are no limits to analyzing this data set; however, some interesting questions to get you started are the following:

- · Is fertility rate significantly associated with life expectancy?
- · Has fertility always been associated with life expectancy?
- Is there a significant difference in life expectancy among the major continent regions? This can be justified first graphically (boxplots, etc.) and next with a hypothesis test.
- Looking at the countries of Sudan, El Salvador, Cape Verde, Bangladesh, etc., consider an association or statistic of your choice. Is the association or statistic different for these countries? How does it change over time?
- Looking at those same countries or the continental regions, create an interactive plot of their demographics using the Google visualization package in R.

Of course, all results need to be justified with the significant statistical hypothesis test. \blacksquare

A Conceptual Approach to Teaching Standard Deviation

By Susan Haller and Melissa Hanzsek-Brill, St. Cloud State University

Students in our mathematics classes usually learn mathematics through a conceptual, hands-on approach: They relate symbolic processes to pictures, models, and vocabulary and learn more than how to simply compute a solution to a problem. Therefore, as we finish the last step in the process for manually calculating the population standard deviation by using an "answer-getting" algorithm, we often find that several students have their hands in the air: "What does it mean?" "Is there a model to help explain those calculations?" Why do we do this?"

The Principles and Standards (NCTM 2000) state that representation facilitates reasoning, supports student understanding, and is the core of mathematical communication. The Common Core State Standards (2010) encourage teachers to use lessons that incorporate the Standards for Mathematical Practice. As longtime mathematics instructors, we were frustrated that, until now, we were unable to provide a conceptual representation of the process involved in calculating standard deviation.

Here, we describe activities to develop a concrete representation of the process used to calculate standard deviation. Because we want to make an explicit connection between this process and the equal sharing method of calculating the mean, we choose to first focus on the calculation of the population standard deviation. We conclude with remarks about how to segue into the calculation of the sample standard deviation.

The authors taught the lesson presented below to different groups of students in introductory mathematics classes, as well as to prospective and practicing middle- and high-school teachers; however, it could be used with students in grades 5–12 and other post-secondary settings as standard deviation is introduced and explored at these levels.

Activity 1: Measures of Central Tendency— What Does the Mean Mean?

The activity begins by asking students what the mean of the data set represents. Students quickly reply that "the mean is what you get when you add all the number together and divide by the number of numbers in the data set." We agree that this is indeed the way to calculate the mean, but that it really doesn't tell us what the mean actually represents. Building a conceptual foundation for the meaning of the mean is critical for building a conceptual foundation for our approach to the standard deviation, so we start with a hands-on approach to this concept.

Students should be assigned to work in groups, and each group gets a container of two different-colored color tiles. In this demonstration, we chose to use red and green tiles. The first data set is {11, 3, 5, 14, 7}, so students are asked to form five



Figure 1

groups of red tiles of size n = 11, 3, 5, 14, and 7. The tiles should be lined up horizontally in an organized manner (Figure 1).

Students are now asked to complete the following tasks using the tiles (see Student Activity 1 Worksheet):

- Push all the red tiles together.
- Redistribute the red tiles so that each of the five line segments has the same number of tiles.
- Compare the steps in each task above with the steps used to calculate the mean of the data set.

Students realize that the concrete model and algorithm contained identical steps. We refer to this approach as the equal sharing method and use it as a springboard to discuss an interpretation of what the mean represents. Students now recognize the mean as the measure of center that 'balances' or 'evens out' the data set.

Students are then engaged in a discussion leading them to understand that the measures of center alone are not sufficient to describe a data set. We next introduce the idea of measures of spread or dispersion, specifically standard deviation for this activity.

Activity 2: The Meaning of the Standard Deviation

Students come to many math classes having had experience calculating the standard deviation either manually or via technology. Some have seen the different formulas used for the sample standard deviation and the population standard deviation; however, they have no conceptual understanding of what standard deviation is. Those who can calculate the standard

9

deviation manually simply memorize the algorithm involved and those who enter the data into a calculator or spreadsheet obtain an answer that does not connect to the mean or the data set, resulting in a weak to nonexistent understanding of this concept.

We have found this lesson works best by beginning with a review of how to manually calculate the population standard deviation, along with a discussion about the sum of the differences from the mean. Note: In classes in which students do not have a preconceived notion of standard deviation, it might be best to do the formula development alongside the modeling. We use the same data set used in the equal sharing method—11, 3, 5, 14, and 7—and begin with finding the mean, then the sum of the differences from the mean (Figure 2). We focus on the population standard deviation, rather than the sample standard deviation, because of the direct link to the equal sharing method of calculating the mean.

Students are often surprised to see that the sum of the differences from the mean is zero. After some discussion, students recognize this sum will always be zero and therefore not useful. We then continue the calculation by squaring each difference from the mean, finding the sum of these squares, and dividing by 5 to find the variance. We finally take the square root of variance to arrive at the calculated population standard deviation of 4 (Figure 3).













Figure 5

Students are asked to examine the given data set and then discuss the meaning of this calculated standard deviation. In short, students don't see any connection between the population standard deviation and the data set. They see that calculating the standard deviation involves differences from the mean, but beyond that, it is basically just a process and another algorithm to memorize. Further, students with a stronger understanding of algebra recognize that taking the square root of the average of squares is not an inverse relationship, causing even more confusion for those who do want a better understanding of the process involved in calculating the standard deviation.

Next, we begin the hands-on approach for modeling the calculation of the standard deviation (see Student Activity 2 Worksheet). Using the same data set as in Activity 1, students are asked to use the green tiles to create line segments of length n = 11, 3, 5, 14, and 7 and then align the red tiles from the end of Activity 1 (that is, the mean of the data set) above or below these segments. By setting up this alignment, we compare the mean of the data set (8) to each number in the original data set. Notice that, in this case, none of the five groups has the same number of green and red tiles—there is always some under- or overlap between the colors (Figure 4).

Students see that many tiles are in one-to-one correspondence, but some are not. We have students remove any tiles in one-toone correspondence (Figure 5), leaving five segments.

The line segments formed by the remaining tiles represent each difference from the mean. We tell students that the length of these segments will be used to represent the length of the side of a square and have students use additional tiles to build each square (Figure 6). For example, if the difference in the number of matched tiles is 3, we add more tiles to build a 3 X 3 square. We refer to these squares as "difference squares." We're quite pleased with the visual this name implies and use it throughout the rest of the activity.



Figure 6



Figure 7

From here, students apply the equal sharing method to find the mean area of the difference squares. They rearrange the tiles in the difference squares to form five equal-area squares (Figure 7).

Students are told that the average area of the difference squares is the variance of the data set. They are reminded that the standard deviation is the square root of the variance and asked what they think standard deviation is in this context and to identify where they see it. They realize that, in this case, the standard deviation is the length of the side of the average difference square.

The final step is to have students relate the modeling of the tiles to the manual calculation of the standard deviation. Students recognize the steps involved in the modeling correspond to the procedure used to calculate the standard deviation. They are asked to 'define' the standard deviation in terms of the squares they constructed and recognize it is the length of one of the sides of the squares they just made.

We next present the students with a second data set: {10, 1, 9, 2, 3, 6, 4, 5}. The mean of this data set is 5, so one of the data values is equivalent to the mean. The challenge in this example is that students must recognize that one of the squares has an area of zero. As students work together, they agree that this area must also be included when finding the average size of the difference squares.

By using their color tiles, students find that the variance of this data set is 9, so the standard deviation—or length of the side of the average difference square—is 3. We follow this hands-on model with the standard algorithm used to calculate the variance and standard deviation and have students compare the steps in the algorithm to the steps in the concrete model.

In the first two data sets presented so far, the variance is a perfect square, so the standard deviation is an integer. One can create an infinite number of data sets with this property by adding a constant k to each data value. This will result in a mean k units greater than that of the original data set and a standard deviation equal to the original data set. We illustrate this in the third example, where 3 is added to each data value in the previous example. This yields the data set {13, 4, 12, 5, 6, 9, 7, 8}, with mean 8 and standard deviation 3. We ask students to compare this example to the previous one, telling them to notice that each data point is now 3 units larger. We ask them to compare the means of the two data sets as well as the areas of their difference squares and ask for an intuitive justification of why the mean should increase by 3 in this case, but the variance and standard deviation should not change. Students already knew the mean would increase by 3, but some are surprised to learn that the variance and standard deviation were unchanged.

We use this as an opportunity to reinforce the concept that variance and standard deviation are measures of spread. If all data points are increased by (in this case) 3, the spread of the data has not changed at all; it has simply shifted 3 units. A logical follow-up question at this point is to ask how



Figure 8

multiplying each data point by three would affect the mean, variance, and standard deviation. Students know the mean should be three times as big and also realize the data are more 'spread out,' so the variance and standard deviation should also be bigger. In this case, the standard deviation is 3 times bigger. Unfortunately, using difference squares to model this does not work well due to the large number of tiles required.

In the fourth example, the data set is {6, 2, 13, 5, 9}, which has a mean of 7 and variance of 14. Although the difference squares could be shared equally among the five groups, 14 is not a perfect square number, so 'perfect' difference squares cannot be constructed using the tiles. Students have to realize that, if a square is to be constructed, its sides will not be of integer length (Figure 8). The manipulative model involving color tiles is starting to 'fall apart,' providing an important lesson as to why the commonly used pencil and paper algorithm is important to develop alongside the concrete model.

Finally, a fifth data set, $\{12, 5, 10, 3, 10, 2\}$, is presented with a mean of 7 and a variance of 88/6 or \approx 14.67. By this time, students recognize that not all difference squares are integers when shared equally. Not only is the standard deviation not an integer, but neither is the variance. We asked students whether they thought this was a serious flaw in the conceptual model. Students agreed that, by the time they got to this example, they already had a good mental image of the conceptual development and that the tiles were no longer essential.

Invariably, at least one student will have seen the formula for calculating the sample standard deviation and want know why we are sometimes told to divide by *n* and other times by *n*-1. We take advantage of this question by reinforcing the differences between population and sample means, as well as sample and population standard deviations. Our process focused exclusively on the population mean and standard deviation. When we have the data for the population, we can calculate the true mean and standard deviation, as was done in the examples presented.

Students understand that, in real life, we don't always know the true population mean and instead know the mean of a sample. Because the sample data will be closer to the sample mean than the true population mean (this idea often takes a fair amount of time for students to absorb), the sum of these 'difference squares' will be smaller than the sum of the 'difference squares' for the population. To correct for this, we need to divide by *n*-1, rather than n. This leads to an interesting discussion about sample size. As n gets larger, the effect of subtracting 1 has less effect than when *n* is small. From this, students begin to recognize less correction is needed as the sample size increases. More specifically, they understand that the sample mean should be closer to the true population mean as *n* increases. Expanding this connection to standard deviation, the sum of the sample 'difference squares' should be closer to the population 'difference squares.'

Conclusion

Understanding standard deviation is challenging, but necessary, when making inferences about data. Too often in our curriculum, the focus is placed on measures of center, which do not always provide a complete picture about a data set. Perhaps this is due to a lack of quality representations related to the process involved in calculating the standard deviation. Both the Principles and Standards and Common Core State Standards for Mathematics emphasize justification of the mathematical processes through modeling and reasoning. It is our hope that the activities presented here will allow students to hold a mental image for the process of calculating standard deviation and connect the algorithm to a concrete model. Our students indicated that visualizing the process through difference squares gave them a better sense of both the steps involved in the calculation process and the impact of the differences from the mean.

References

- Common Core State Standards Initiative (CCSSI). 2010. Common Core State Standards for Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers. www.corestandards.org/assets/CCSSI_ Math%20Standards.pdf
- National Council of Teachers of Mathematics. 2000. *Principles and Standards for School Mathematics*. Reston, VA: NCTM.



Teaching Statistics Through Data Investigations

www.mooc-ed.org/tsdi

Learn with colleagues near and far in this 8-week, online professional development course, designed for teachers of statistics in grades 6-12 and post-secondary contexts. This course can help you learn to teach statistics using investigations, with real data and real cool tools!

The course is **FREE** and can lead to continuing education credits.

Offered through the Friday Institute for Educational Innovation at NC State University Lead instructor: **Dr. Hollylynne Lee**, Professor of Mathematics Education



ASA PR Campaign for Students, Teachers, Counselors Continues Expanding Reach

More "This is Statistics" Material to Share with Your Students, Colleagues

This is Statistics (*http://thisisstatistics.org*), the American Statistical Association's public relations campaign aimed at making students aware of the benefits of taking THIS a statistics course (or courses) and the opportunities a career in statistics offers, continues to increase its reach to students and those around students (e.g., teachers, counselors, parents). Since its launch at the 2014 STATISTICS Joint Statistical Meetings, the This is Statistics (TiS) website has averaged nearly 5,000 unique visitors per month and its six videos (www.youtube.com/user/ ThisIsStats) have been watched nearly 66,000 times. The number of likes on its Facebook page is 6,400, and the Twitter account has 3,100 followers. Despite these encouraging numbers, we need your help, as much work remains to be done to reach students and those who help influence their educational and career decisions.

As a short update on TiS, new profiles, news stories, and videos were added over the last several months (see New Resources). The profiles include statisticians and data scientists, including Rayid Ghani, the Obama campaign's data chief. Recent news stories center on the demand for statisticians and the growing number of students taking statistics courses and/or seeking statistics degrees. Several videos were also released, including videos produced by ASA student chapters and a professionally produced video featuring Megan Price of the Human Rights Data Analysis Group. The communications firm executing the

TiS campaign—Stanton Communications—also successfully pitched many stories to the media about the tremendous growth in statistics degrees and the demand for statisticians.

New Resources

Profiles

Detroit Fire Department Statistician Cassie DeWitt http://bit.ly/1QyEr4d

Georgetown Professor Kimberly Sellers *http://bit.ly/1SQGtfg*

Obama Campaign Data Chief Rayid Ghani http://bit.ly/1URuAtM

Pegged Software Data Scientist Shannon Cebron http://bit.ly/10VAPY6

Quora Has Questions, Data Scientist Olivia Angiuli Has Answers http://bit.ly/1Sr8hYg

Hilary Parker Gets Crafty with Statistics in Her Not-So-Standard Job *http://bit. ly/1UXknum* Video on Megan Price of the Human Rights Data Analysis Group (https://youtu.be/orWo1w8a4zY)

News stories/Features/ Press Releases

http://thisisstatistics.org/latest-news

Florence Nightingale: The Lady with the Data http://bit.ly/1Uxe9Dj

Wharton Statistics Program 'Bursting at the Seams' http://bit.ly/21SmmTS

Seattle Times: Statistics Careers Are Hot http://bit.ly/1TDu1zB

More Students Earning Statistics Degrees, but Not Enough to Meet Surging Demand for Statisticians http://bit.ly/1IXT8HH What's the Best Job in Business? Statistician, According to U.S. News & World Report http://bit.ly/1VQ5Xgm

Data Scientist Most In-Demand Job, Reports Fast Company http://bit.ly/1Sr8J8Y

Statistician Projected as Top 10 Fastest-Growing Job http://bit.ly/1RFD6Fs

Data-Driven Government Means Greater Role for Statisticians, Says White House Chief Data Scientist http://bit.ly/1Rwu12E For 2016, in addition to keeping the website fresh through new profiles and news items, we are undertaking fresh activities,



including more direct outreach to teachers and counselors. 2016 ASA President Jessica Utts also has an initiative—"Statistics Careers for AP Statistics and Other K–12 Classrooms," chaired by ASA Board member Anna Nevius—that is already helping to advise about how to better reach highschool students.

The ASA asks you to share the TiS resources and URL with your networks, especially students, parents, teachers, and counselors. See

"How You Can Help" for more ways to get the word out on statistics. ■

How You Can Help

Share This is Statistics with your students, fellow teachers, counselors, and others:

- Post the TiS logo to your website
- Follow TiS on Twitter (@ThisisStats)
- Like the Facebook page: *www.facebook.com/ThisisStats*
- Share TiS resources through your networks, including on social media
- Use the TiS Promotional Toolkit: http://thisisstatistics.org/educators
- Post/distribute its one- and two-page posters
- Consult talking point PowerPoint slides



This special offer is tailored for K–12 educators so you can enhance your students' statistical education.

Activate your free trial membership at www.amstat.org/kl2trial.

Free trial membership is valid for new ASA members only.

Members receive:

Access to professional learning community teaching resources, including webinars, peer-reviewed lesson plans (STEW), Census at School, and publications such as the Statistics Teacher Network, GAISE: A Pre-K–12 Curriculum Framework, Bridging the Gap Between Common Core State Standards and Teaching Statistics, and Making Sense of Statistical Studies.

Information about upcoming events and services for K–12 teachers and students, including workshops, student competitions, data sources, and publications.

Access to the **ASA Community's K–12 discussion group**, where like minds share ideas, questions, and resources.

Subscriptions to *Amstat News*, the ASA's monthly magazine, and *Significance*, a magazine aimed at international outreach and statistical understanding.

Members-only access to the ASA's top journals and resources, including online access to CHANCE magazine, the Journal of Statistics Education, and The American Statistician. Additional resources accompanying this lesson also are posted.

Odd or Even? The Addition and Complement Principles of Probability



Carrie Even Waterloo Community School District

Susannah Weaver Central College weavers1@central.edu STatistics Education Web

online journal of K-12 statistics lesson plans

Published: November, 2015

evenc@waterlooschools.org

Overview of Lesson

After an exploration activity and a class discussion, students will be able to determine when events are mutually exclusive, and then use the Addition Principle to calculate the probability of mutually exclusive and non-mutually exclusive events. Furthermore, the students will explore the probabilities of complementary events and develop the Complement Principle.

GAISE Components

This investigation follows the four components of statistical problem solving put forth in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*. The four components are: formulate a question, design and implement a plan to collect data, analyze the data, and interpret results in the context of the original question.

This is a GAISE Level B activity.

Common Core State Standards for Mathematical Practice

MP1. Make sense of problems and persevere in solving them. MP3. Construct viable arguments and critique the reasoning of others.

Learning Objectives Alignment with Common Core State Standards and NCTM Principles and Standards for School Mathematics

Learning Objectives	Common Core	NCTM PSSM
Students will use simulations to	7.SP.C.8.C. Design and	Grades 6-8 Understand and
explore empirical probabilities and	use a simulation to	apply basic concepts of
compare them to theoretical	generate frequencies for	probability: use proportionality
probabilities.	compound events.	and a basic understanding of
		probability to make and test
		conjectures about the results of
		experiments and simulations;
		Grades 9-12 Understand and
		apply basic concepts of
		probability: use simulations to

		construct empirical probability
		distributions:
Students will determine sample spaces using organized tables. Students will use sample spaces to compute probabilities of compound events using unions and complements.	 7.SP.C.8.B. Represent sample spaces for compound events using methods such as organized lists, tables, and tree diagrams. S-CP.1. Describe events as subsets of a sample space, using characteristics of the outcomes or as unions, intersections, or 	distributions; Grades 6-8 Understand and apply basic concepts of probability: compute probabilities for simple compound events, using such methods as organized lists, tree diagrams, and area models. Grades 9-12 Understand and apply basic concepts of probability: understand the concepts of sample space and probability distribution and construct sample spaces and
	complements of other	distributions in simple cases.
	events.	
Students will discover and use the Addition Principle for mutually exclusive and non-mutually exclusive events: If A and B are mutually exclusive, then P(A or B) = P(A \bigcup B) =P(A) + P(B). If A and B are not mutually exclusive, then P(A or B) = P(A \bigcup B) = P(A) + P(B) – P(A and B).	S-CP.7. Apply the Addition Rule, $P(A \text{ or } B) = P(A) + P(B) - P(A and B) and interpret theanswer in terms of themodel.$	Grades 6-8 Understand and apply basic concepts of probability: understand and use appropriate terminology to describe complementary and mutually exclusive events;
Students will discover and use the complement principle: $P(A^c) = 1 - P(A)$	S-CP.1. Describe events as subsets of a sample space, using characteristics of the outcomes or as unions, intersections, or complements of other	Grades 6-8 Understand and apply basic concepts of probability: understand and use appropriate terminology to describe complementary and mutually exclusive events:

Prerequisites

Students will need to be familiar with finding the probability of a single event, A: $P(A) = \frac{Number of Outcomes for Event A}{Number of Total Possible Outcomes}$ Students will be familiar with the vocabulary associated with probability in the table below and have experience determining sample spaces and subsets. Students will be familiar with Venn diagrams. Needed Ve 1 1

Needed Vocabulary			
Subset	A set that is part of a larger set.		
Theoretical Probability	The number of ways an event can occur divided by the total number of possible outcomes.		
Experimental Probability	Is based on repeated trials of an experiment or (# of Successes) / (Total # of Trials).		

Time Required

This lesson will take place in two 50-minute lessons.

Materials and Preparation Required

- Student Worksheet
- TI Nspire Calculator or Core Math Tools
- Instructions for TI-Nspire Calculator Die Simulator or Core Math Tools (see Appendix)
- Two standard numbered cubes for each pair of students

Rationale for Lesson design:

Students find it difficult to find the probability of compound events (Carpenter 1981). In a study of compound events where students summed up pairs of numbers, Fischbein et al. (1991) found that students do not possess any natural intuition when it comes to finding probabilities of compound events. Students may possess a tendency to globally evaluate the size of the sample space and its structure, and they do recognize the relationship between probability and the size of the corresponding sample space. From this they can develop a natural, intuitive tendency to evaluate the probability based on the sample space. Thus, evaluating the probability of compound events requires an understanding of the sample space (a pre-requisite that is revisited in this activity).

Shaughnessy (1977) found that conventional lecture may not be the best way to overcome student misconceptions of probability (such as the availability and representativeness heuristics presented by Kahneman & Tversky (1974)). Instead, he showed that student misconceptions could be addressed through hands-on experiments and activities in which students discovered probability concepts for themselves. He also showed that misconceptions can be discovered by having students compare conjectures to results obtained via experiments (1981). Furthermore, Shaughnessy and Zawojewski (1999) found that students need opportunities such as that presented in this lesson plan to work through substantive probability problems.

Finally, Beitzel, et al. (2011) found that students using only concrete representations such as Venn Diagrams to solve probability problems involving non-mutually exclusive events were outperformed by students instructed in using only the equation for the addition rule. Students that had to construct the Venn diagram to solve probability problems reported a higher cognitive demand than those using just the abstract representation of the formula. Thus, it is critical that students develop an understanding of the addition rules in order to be able to use them without constructing a Venn diagram.

The main author was unable to find any research on what students understand about the complement rule for probability nor how to teach this concept. As a result, we followed the idea of exploring Venn diagrams and discovering the Complement Principle by looking at subspaces.

Odd or Even? The Addition and Complement Principles of Probability Teacher's Lesson Plan

Describe the Context and Formulate a Question

Two people are playing a game of "Odd or Even" with two standard number cubes. (Each student should have the **Student Handout** in the Appendix). A <u>trial</u> occurs when two cubes are rolled once and points are assigned. The <u>game</u> consists of many trials, the number of which are decided upon by Player A and B before the first trial. Consider the following rules.

Odd or Even Rules:

Roll two standard number cubes. If the sum is 6 or an odd number, then Player A scores a point. If the sum is 6 or an even number, then Player B scores a point. Repeat the above steps an agreed number of times. Whomever has the most points at the end of the game wins.

Ask students: "Suppose you play a game consisting of many trials. Predict which partner is likely to score more points in "Odd or Even". Justify your reasoning.

Students will begin by making a prediction for the game "Odd or Even". As students often possess misconceptions about probability, elicit responses to question 1 on the Student Handout to determine whether misconceptions about sample size exist, as developing intuition about events relative to sample size is often a misconception (Fischbein 1991). A common misconception is that a student is either Player A or Player B. Be sure to address this with students by emphasizing that we are analyzing the game, not playing it as the two players.

Collect Data

Next, show students two six-sided number cubes and ask how they might determine who would win in a game of 1000 trials without actually playing 1000 times. To help address students' misapplied confidence in small numbers, we will first look at samples of size 30 (Hope 1983). Before having students play the game (neither partner is Player A nor B), ask them what assumptions we are making about the game. Students should determine that we are assuming the number cubes are fair and that each side has the same probability of landing face up (we have a uniform probability distribution of the sides of the cube).

After students complete a game of 30 trials, bring the class together and ask what they found out. Begin to discuss whether 30 trials is enough. Did everyone get the same probabilities? Which ones are more accurate? The instructor may wish to pool the class data and ask whether students feel more confident about those results. Be sure to talk about underlying assumptions such as whether everyone tossed the cubes the same way. With the pooled data, students may still believe that we have enough trials to be fairly certain. How certain do we want to be? What if we wanted to predict the winner in a game consisting of 1000 trails? How could we do that without

actually playing 1000 times? Lead students to discover that it would be great if we could automate the process and use technology to perform many trials.

Question 4. Students will next simulate 1000 trials of the game "Odd or Even" using the TI-Nspire calculator or Math Core Tools (see Appendix A for directions on how to use both tools). After the simulation, have students compare their results with their original prediction (question 1) and their sample of size 30 (question 2). Give students the opportunity to change their original conjecture with an explanation. The process of moving from 30 trials to a class pooled data set to a simulation of 1000 trials, will help students combat the misconception of small numbers.

By the time students get to question 5 on the Student Handout, most will be skeptical that the game is fair. What does it mean to be fair? Which player would you rather be?

Before moving on, discuss/define the mathematical term "mutually exclusive" (two events are mutually exclusive if they have no outcomes in common). To help solidify the definition, ask students to describe some events in their everyday life that are mutually exclusive.

In the first part of the Explore section, have the students look at the three Venn diagrams. Before moving on, ask why two are labeled "not mutually exclusive" and one is labeled "mutually exclusive". Then, discuss their similarities and differences.

Analyze Data

The students will then compute the probabilities of two events that are mutually exclusive using simple probability and sample spaces. Next, students are asked to think about the relationships between P(A), P(B) and P(A or B). Students should discover that the sum of two probabilities is the third, and conclude that when A and B are mutually exclusive events, P(A or B) = P(A) + P(B).

Upon completing questions 7 - 11 on the Student Handout, make sure the students discuss their solutions first with their partner and then share out as a class. When sharing out as a class, define the addition rule of mutually exclusive events mathematically and write the definition in words the students understand. Sharing first with each other in small groups (in their native language, if necessary), is most important for the ELL students.

For Questions 12 - 17, students will find the probability of two events that are not mutually exclusive by using simple probability. They will then compare how they computed the probability of Player A winning to how they computed the probability of Player B winning. As part of this process, students will compare and contrast the results of mutually exclusive and not mutually exclusive events; they will conclude that if events are not mutually exclusive, then P(A or B) = P(A \cup B) = P(A) + P(B) - P(A and B). After completing question 12, have students share with their partner and then as a class. Upon completing questions 13 - 17, bring the class together to discuss what they discovered. When they share out as a class, define the addition rule for the probability of two non-mutually exclusive events mathematically and write it in words the students understand. Sharing first with each other in small groups (in their native language, if necessary), is most important for the ELL students.

Question 18 asks students to return to the original question – who would win in a game consisting of many trials? Bring students together as a class and have them compare the three probabilities that they have found: the one using dice and 30 trials of the game, the simulation of 1000 trials, and the theoretical probability. They should conclude that the simulation of 1000 trials of the game gave the most accurate experimental probability.

Once the addition rules are derived, the students will find the probability of the complement of Player A winning. A common misconception is that the complement of Player A winning must be player B winning. The way the game is designed, neither player may win: they could tie. They will then compute 1 - P(A) and compare this result with the one found using simple probability. The students will then define the probability of the complement of an event A: $P(A^c) = 1 - P(A)$. Upon completing question 19 again have students discuss at their tables and share out as a class. Both ELL students and non- ELL students benefit by bouncing ideas off each other, and the discussion helps identify misconceptions. The instructor may choose to address misconceptions at the table-level or, if a misconception is prevalent throughout, address it as a class. Sometimes even if the misconception is not prevalent, it should be used for good classroom discussion to address non-examples.

Interpret Results

Key points of the exploration that need to be discussed as a class are questions 11g, 17f, and 19g on the Student Handout. Following these key questions, the students will write in words and corresponding mathematical notation the two versions of the addition rule for probability and define the complement in the boxed areas. For ELL students, have them write in their native language or words that make sense to them along with the mathematical notation. It is vital that they understand the definitions for the Addition Principle for mutually exclusive and non-mutually exclusive events and the Complement Principle.

Question 11g. If two events, M and N, are mutually exclusive, then P(M or N) = P(M) + P(N); $P(M \cup N) = P(M) + P(N)$; Since M and N have nothing in common, or do not intersect, the events are mutually exclusive so you add the two probabilities together. "OR" is used to show that a sum may be in either A or B or in both, and is defined as the union that is notated by "U" symbol.

Question 17f. If two events, H and I, are non-mutually exclusive, then P(H or I) = P(H) + P(I) - P(H and I); $P(H \cup I) = P(H) + P(I) - P(H \cap I)$; Since H and I overlap they are non-mutually exclusive, and therefore we add the probabilities and subtract the probability of what they have in common (or the overlapping part) because it is included twice. "AND" is used to show the intersection of the two sets and notated by " \cap " symbol.

Question 19g. If Q and R are complementary events, then $P(R) = P(Q^c) = 1 - P(Q)$; Find the probability of the complement of Q, which is the probability of not Q and is written as $P(Q^c)$. Since 1 is the total probability of an event you can subtract a part, P(Q), to obtain the remaining probability, $P(Q^c)$.

The final two questions ask students to reflect on the definitions (mutually exclusive, complementary) in relation to the Player A and Player B winning. They are not mutually

exclusive and they are not complementary. Students should be given time to think about these answers before developing their own game. A common misconception will be that either Player A or Player B wins, however, because they both score a point when the sum is 6, the game could end in a tie if an even number of rounds is played. In this case, neither player has more points than the other and neither wins.

Suggested Assessment

- 1. A single letter is chosen at random from the word MATHEMATICALLY. Let A = "the letter E" and B = "vowel". Are events A and B mutually exclusive? Explain, why or why not.
- 2. A standard number cube is rolled. Let event E be "an even number is rolled" and event O = "an odd number is rolled". Are events A and B complementary? Why or Why not?
- 3. If the probability of Jack winning the Dinosaur Explosion game is $\frac{1}{3}$ and the probability of Leo winning is $\frac{1}{5}$, what is the probability that either Jack or Leo will win the game? Support your answer with work.
- Solve using two different methods. Make sure you show your work and explain the differences between your methods.
 The integers 1, 2, 3, ..., 20 are written on slips of paper and placed in a bowl and thoroughly mixed. A slip is drawn from the bowl at random. What is the probability that the number on

the slip is either prime or divisible by 3?

Method 1:

Method 2:

Explanation of the differences of method 1 and method 2:

- 5. Suppose F is the event that you draw a four and C is the event that you draw a club. Suppose that, $P(F) = \frac{4}{9}$, $P(C) = \frac{4}{9}$, and $P(F \cap C) = \frac{1}{3}$. Find probability that you draw a four or a club.
- On April 15,1912, the Titanic struck an iceberg and rapidly sank with only 710 of her 2,204 passengers and crew surviving. Data on survival of passengers (<u>http://www.encyclopediatitanica.org/titanic-statistics.html</u>) are summarized in the table below.

	Survived	Did not survive	Total
1st Class Passengers	201	123	324
2nd Class Passengers	118	166	284

3rd Class Passengers	181	528	709
Total Passengers	500	817	1317

- a. What is the probability that someone on the Titanic was a 2^{nd} or a 3^{rd} class passenger?
- b. What is the probability that someone on the Titanic was a 2nd class passenger or survived?

Assessment Answers:

- 1. No, A and B are not mutually exclusive because E is a vowel and therefore A and B intersect.
- 2. Yes, E and O are complementary because their union is the whole sample space and the events are disjoint. P(S) = 1, and 1 = P(E) + P(O)
- 3. $\frac{1}{3} + \frac{1}{5} = \frac{8}{15}$
- 4. Method 1: Counting principle or simple probability

List all the numbers: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 Count the numbers that are either prime or divisible by 3 without duplication 14/20= 7/10

Method 2: Addition Rule for non-mutually exclusive events.

P(Prime or Divisible by 3)

 $P(Prime) + P(Divisible by 3) - P(Prime \cap Divisible by 3) = 9/20 + 6/20 - 1/20 = 14/20 = 14/20$ 7/10

Method 3: Complements

You count the evens and take out the number 2 because it is a prime and then take out the numbers that are divisible by 3. That leaves you with probability of 6/20. Therefore 1 - 6/20 = 14/20 = 7/10

Possible explanation of the differences of methods:

The first method used the counting principle and the second method used the addition principle. With the counting principle, one counts the numbers in the set and with the addition principle, one finds the probability of a prime, the probability of a number divisible by three and the probability that a number is divisible by three and a prime. The addition rule is used to calculate the probability.

5.
$$\frac{4}{3} + \frac{4}{3} - \frac{1}{3} = \frac{5}{3}$$

5. $\frac{1}{9} + \frac{1}{9} - \frac{1}{3} = \frac{1}{9}$ 6. a. $P(2^{nd} \text{ class or } 3^{rd} \text{ class}) = P(2^{nd} \text{ class}) + P(3^{rd} \text{ class}) = \frac{284}{1317} + \frac{709}{1317} = \frac{993}{1317} = \frac{331}{439}$ b. $P(2^{nd} \text{ class or } \text{Survived}) = P(2^{nd} \text{ class}) + P(\text{Survived}) - P(2^{nd} \text{ class and } \text{survived}) = \frac{284}{1317} + \frac{500}{1317} - \frac{166}{1317} = \frac{618}{1317} = \frac{206}{439}$

Possible Differentiation

To extend this activity, consider looking at probabilities involving variables and geometric probabilities as in the following questions.

Extension

- 1. Suppose H is the event that you buy a two-story house and T is the event tan houses.
 - a. Suppose $P(H) = \frac{2}{3}$, $P(T) = \frac{3}{5}$, and $P(H \cup T) = \frac{13}{15}$. Find the probability that you buy a tan two-story house.
 - b. Suppose $P(H) = \frac{1}{2}$, $P(H \cap T) = \frac{1}{4}$, and $P(H \cup T) = \frac{2}{3}$. Find the probability that the house is tan.
 - c. Suppose P(H), P(T), and P(H \cap T) can be represented by the following expressions respectively $\frac{2}{x}$, $\frac{3}{2x}$, and $\frac{1}{x}$. Find the probability that you buy a house that is two-story or tan.
- 2. What is the probability of scoring exactly 30 points with one dart thrown? The bonus region triples your score and each ring is 4 in. thick. The central angle for the bonus sector is 30° .

Extension Solutions

1.

a.
$$P(H \cap T) = \frac{5}{5}$$

b. $P(T) = \frac{5}{12}$
c. $P(H \cup T) = \frac{5}{23}$

2. Let A = "Area of 30 without Bonus" Area of 30 without Bonus = $\frac{300}{360}(4^2\pi) = \frac{40}{3}\pi$ P(A) = $\frac{\frac{40}{3}\pi}{256\pi} = \frac{5}{96} = .05208$

> Let B = "Area of 10 with Bonus" Area of 10 with Bonus = $\frac{60}{360}(12^2\pi - 8^2\pi) = \frac{8}{3}\pi$ P(B) = $\frac{\frac{8}{3}\pi}{256\pi} = \frac{1}{96} = .0104$

$$P(A \text{ or } B) = P(A) + P(B) = \frac{6}{96} \text{ or } .0625$$



References

Beitzel, Brian D., Richard K Staley, and Nelson F DuBois. (2011). When Best Intentions Go Awry: The Failures of Concrete Representations to Help Solve Probability Word Problems. *Educational Research Quarterly Vol. 34.3*, 3-14.

Carpenter, Thomas P., Mary Kay Corbitt, Henry S. Kepner, Jr., Mary Montgomery Lindquist and Robert E. Reys. (1981). What are the Chances of your Students Knowing Probability? *The Mathematics Teacher* 74, 342-344.

Fischbein, Efraim, Maria Sainati Nello and Maria Sciolis Marino. (1991). Factors Affecting Probabilistic Judgements in Children and Adolescents. *Educational Studies in Mathematics, Vol* 22,(6), 523-549.

Hope, Jack A. and Ivan W. Kelly. (1983) "Common Difficulties with Probabilistic Thinking." *The Mathematics Teacher. Vol.* 76 (8), 565-570.

Kahneman, Daniel and Amos Tversky. (1974) Judgment under Uncertainty: Heuristics and Biases. *Science, New Series, Vol. 185* (4157), 1124-1131.

Mosteller, Frederick, Robert E. K. Rourke, George Brinton Thomas. (1970). *Probability, a First Course* (p 84). Addison-Wesley Publishing Co.

Shaughnessy, J. Michael. (1977) Misconceptions of Probability: An Experiment with a Small-Group, Activity-Based, Model-Building Approach to Introductory Probability at the College Level. *Educational Studies on Mathematics, Vol.* 8 (3), 285-316.

Shaughnessy, J. Michael and Judith S. Zawojewski. (1999). Secondary Students' Performance on Data and Chance in the 1996 NAEP. *The Mathematics Teacher Vol. 92* (8), 713-718.

Shaughnessy, J, Michael. (1981). Misconceptions of Probability: From Systematic Errors to Systematic Experiments and Decisions. In *Teaching Statistics and Probability* (pp 90-100). National Council of Teachers of Mathematics.

Further Reading About the Topic

Kelley, I. W. and F.W. Ziers. (1996). Mutually Exclusive and Independence: Unraveling the Basic Misconceptions in Probability Theory. Retrieved from https://www.stat.auckland.ac.nz/~iase/publications/icots2/Kelly.Zwiers.pdf.

Hirsch, Linda S. and Angela M. O'Donnell. (2001). Representativeness in Statistical Reasoning: Identifying and Assessing Misconceptions. Retrieved from http://www.amstat.org/publications/jse/v9n2/hirsch.html.

Odd or Even? The Addition and Complement Principles of Probability Student Handouts

Launch:

Work with a partner to analyze the game "Odd or Even". A <u>trial</u> occurs when two cubes are rolled once and points are assigned. The <u>game</u> consists of many trials, the number of which are decided upon by Player A and B before the first trial.

Odd or Even Rules:

Roll two standard number cubes. If the sum is 6 or an odd number, then Player A scores a point. If the sum is 6 or an even number, then Player B scores a point. Repeat the above steps an agreed number of times. Whomever has the most points at the end of the game wins.

- 1. Suppose you play a game with a certain number of trials. Predict which partner is likely to score more points in "Odd or Even". Justify your reasoning.
- 2. Next, ask your teacher for two standard numbered cubes. Then, play the game with 30 trials, keeping track of scores by tallying points:

Player A's points:	Probability that Player A wins:
Player B's points:	Probability that Player B wins:

3. Looking at your results, which player was more likely to win?

How do your experimental results compare with your prediction in question 1? Are your results surprising? In what way(s)?

4. Using the technology specified by your teacher, simulate tossing a pair of numbered cubes 1000 times. How many points did each player win?

Player A's points:	Probability that Player A wins:
Player B's points:	Probability that Player B wins:

How do your results using technology to simulate playing the game compare with the prediction you made in question 1 and your experimental result in question 3?

- 5. Do you think the game is fair? Why or why not?
- 6. We want to be sure of our answer to the previous question. What ideas to you have to find the exact/theoretical probability of each player winning?

Explore: On Your Own...



7. Consider the Venn Diagrams above where one circle in each diagram represents S = "the sum is 6" and the other represents O = "the sum is odd".

a. The Venn diagram that represents Player A's events is diagram ______. Explain why.

b. How is the probability of Player A winning related to the diagram you chose in the previous part?

With your Partner...

- 8. The table to the right represents the sample space for the sum of the values when two cubes are tossed. Complete the table.
- 9. Using the table to the right, shade the cells that represent Player A winning the game.
- 10. What is the probability that Player A wins the game? That is, what is the probability of rolling a sum of 6 or a sum that is odd?

P(Player A wins) = P(S or O) =

11. For each part, shade the corresponding event and use it to calculate the probability. a. Shade event "S", and find P(S). b. Shade event "O", and find P(O).

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

c. Compare the diagrams for "S" and "O" in the previous question with the diagram for "S or O" in question 9. What, if anything do you notice?

d. How do you think P(S), P(O), and P(S or O) are related? Explain your answer.

e. Test your conjecture to part d. by computing probabilities. Were you correct or do you want to change your conjecture?

f. Why do you think the relationship you found in part e. is true? Think about the event and sample spaces and the Venn diagram.

g. Summarize If two events, M and N, are ______, then $P(M \text{ or } N) = P(M \cup N) =$ _____.

If two events, M and N, are	, then
P(M or N) =	Memorize
	Me

On Your Own...

- 12. Now suppose we re-label the circles in the Venn Diagrams above so that one circle in each diagram represents S = "the sum is 6" and the other represents E = "the sum is even".
 - a. The Venn diagram before question 7 that represents Player B's events is diagram _____. Explain why.

b. Redraw the Venn diagram from the previous part. Shade and describe in your own words the region where S and E overlap.



c. How is the probability of Player B winning related to the diagram you chose in the previous part?

With your Partner...

- 13. Using the table to the right, shade the cells that represent Player B winning the game.
- 14. What is the probability that Player B wins? That is, calculate the probability of rolling a sum of 6 or a sum that is even.

P(S or E) =

15. Create two sample spaces and represent event "S" and event "E".

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

16. What are P(S) and P(E)?

$$P(S) = P(E) =$$

17. Compare the diagrams of your sample spaces for the events "S" and "E" with that of "S or E".a. What do you notice?

b. Explain your discovery using the Venn diagram from question 12.

c. How do you think P(S), P(E), and P(S or E) are related? Explain your answer.

d. How is your answer in part c related to your answer from question 11d?

e. What is the probability that Player B wins?

P(Player B wins) = P(S or E) =

f. Summarize	
If two events, H and I, are	 ,
then $P(H \text{ or } I) = P(H \cup I) =$	 <u>.</u>

If two events, H and I, are	
then P(H or I) =	Memorize Me!

18. We found the theoretical probability that Player A wins one point and the probability that Player B wins one point. Return to the original question: if you play a game with many trials, who is likely to win? Explain your answer.

On Your Own...

19. a. Shade the area of the Venn diagram to the right that represents the probability of not getting either a sum of 6 or an odd sum.

b. Using the sample space table to the right, calculate the probability of getting a sum that is neither 6 nor odd using simple probability.

 $P([S \cup O]^{c}) =$

With your Partner...

c. Compare your sample space diagram for $S \cup O$ (question

9) with the sample space diagram for $[S \cup O]^c$ (question 19a). Compare also the Venn diagram for $S \cup O$ (question 7) with the Venn diagram for $[S \cup O]^c$ (question 19b). What do you notice?

d. How do you think the probabilities $P(S \cup O)$ and $P([S \cup O]^c)$ are related? Explain your answer.

f. Use $P(S \cup O)$ to compute $P([S \cup O]^c)$.

g. Summarize

Two events A and B are called *complementary* if A and B are mutually exclusive events whose union is the universal space. Symbolically, we write $A = B^{c}$ and $B = A^{c}$. If S and T are complementary events, then $P(T) = P(S^{c}) =$ _____.



20. Are the events of Player A winning a point and Player B winning a point mutually exclusive? Complementary? Support your answers using theoretical probabilities computed in previous questions.

Close:

21. Now it's your turn. Design a game using two six-sided number cubes so that each of Player A and Player B can win in one of two ways, but the two players do not have the same probability of winning. Calculate the probability that each player wins.



Mutually Exclusive

wintually Exclusive									
+	1	2	3	3 4 5					
1	2	3	4	5	6	7			
2	3	4	5	6	7	8			
3	4	5	6	7	8	9			
4	5	6	7	8	9	10			
5	6	7	8	9	10	11			
6	7	8	9	10	11	12			

Student Handout Solutions for Teacher

- 1. Partner A because the two events have nothing in common.
- 2. Answers will vary.
- 3. Answers will vary, but it is likely that most pairs of students determined that Player A won.
- 4. Answers will vary, but it is likely that with 1000 trials most pairs of students determine that Player A won.
- 5. Not fair because the outcomes will not be the same for partners A and B. A will have more because 6 is never an odd number and therefore the two conditions never intersect. Player B will have less because 6 is included as an even number.
- 6. Answers will vary.

7.

- a. Diagram Y, because 6 will never be an odd number so the two events cannot overlap.
- b. Because S and O do not intersect I can conclude that there should be more outcomes because of no overlap and that is why partner A should win. Partner A has a greater probability than Partner B.
- 8. Table shown to the right.
- 9. Table shown to the right.

 $10.\frac{23}{36}$

11.

4 5 $\frac{9}{10}$ P(S) = $\frac{5}{36}$ $\frac{6}{9}$ P(O) = $\frac{18}{36}$

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

b.

a.

- c. The diagrams for "S" and "O" combine to give the diagram for "S or O"
- d. It appears that P(S) + P(O) = P(S or O) because the event spaces for S and O don't have any overlap.
- e. $P(S) + P(O) = \frac{5}{36} + \frac{18}{36} = \frac{23}{36}$ while $P(S \text{ or } O) = \frac{23}{36}$. They are equal, so our conjecture was correct.
- f. Because the event spaces do not intersect, their circles do not intersect in the Venn diagrams or are mutually exclusive.

g. If two events, M and N, are <u>mutually exclusive</u>, then $P(M \text{ or } N) = P(M \cup N) = \underline{P(M)} + \underline{P(N)}$.

12.

- a. Diagram Z, because the number 6 is also an even number, therefore you know that it is contained in both sets S and E causing the overlapping of sets
- b. Since the sum of 6 is a sum that is an even number, then the event S is a subset of the event E, and therefore lies entirely within set E.
- c. Because of the intersection of S and E you cannot take the 6 into account twice. Therefore B's probability of winning should be smaller.

13.

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

14. P(S or E) =
$$\frac{18}{36}$$

15.

+	1	2	3	4	5	6	
1	2	3	4	5	6	7	
2	3	4	5	6	7	8	
3	4	5	6	7	8	9	Event E,
4	5	6	7	8	9	10	
5	6	7	8	9	10	11	
6	7	8	9	10	11	12	
1	1	2	2	4	5	(

16. $P(E) = \frac{18}{36}$ and $P(S) = \frac{5}{36}$

17.

- a. The two events overlap, in fact, event S is a subset of event E.
- b. We chose Diagram Z because S is a subset of E.
- c. We can't simply add the probabilities because we would be adding P(S) twice, once when we add P(E) and once when we add P(S).
- d. We need to subtract the intersection when events are not mutually exclusive: P(E) + P(S) P(S and E) = P(E) since P(S and E) = P(S)

- e. P(Player B wins) = P(S or E) = $\frac{18}{36}$
- f. If two events, H and I, are not mutually exclusive, then $P(H \text{ or } I) = P(H \cup I) = \underline{P(H)} + \underline{P(H)$ $P(I) - P(H \cap I)$.

18. In 100 games, Player A is expected to win 100 * $\frac{23}{36} = 63.9$ points while Player B is expected to win 100 * $\frac{18}{36}$ = 50 points. Thus, Player A is more likely to win.

19.

- a. The region outside both circles is shaded.
- b. $P([S \cup O]^c) = \frac{13}{36}$
- c. In both the sample space diagram and the Venn diagram, the events have no overlap and they make up the whole space.
- d. The probabilities add to one.
- e.
- e. $\frac{23}{36}$ f. $1 \frac{23}{36} = \frac{36}{36} \frac{23}{36} = \frac{13}{36}$ g. If Q and R are complementary events, then P(R) = 1 P(Q).
- 20. The events that Player A wins and that Player B wins are not mutually exclusive (they overlap because both can win a point if the sum is 6) and they are not complementary because they are not mutually exclusive (though their union is the sample space).
- 21. Answers will vary.

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Appendix

TI-Nspire and TI-Nspire CAS Simulation Directions

Directions can also be found at

http://math.kendallhunt.com/documents/daa2/cntns/daa2cntns_014_10.pdf

To simulate tossing two standard numbered cubes, we need to assume that each number has an equally likely chance of occurring. In other words, we need a uniform distribution of the numbers 1 through 6 so that we can use the random number generator in the List & Spreadsheet application.

- 1. Open the List & Spreadsheet app.
- 2. We have two numbered cubes and want to find their sum, so label columns A, B, and C "cube1", "cube2", "sumcube", respectively.
- 3. To simulate rolling a six-sided cube 1000 times, type "randint(1, 6, 1000)" into the formula cell of columns "cube1" and "cube2".
- 4. To compute the sum of the two cubes, in the formula cell for "sumcube" type "cube1 + cube2". Then, press "Enter" to populate the column.
- 5. To see the distribution of the sums, you can create a histogram: from the Home screen, go to **Data & Statistics**.
- 6. Add "sumcube" to the x-axis. Select "Menu" and "PlotType | Histogram"
- 7. From the Histogram you can tally the number of outcomes that make up Player A's and B's outcome.

Core Math Tools Simulation Instructions:

Another way of making a simulation if you didn't have access to the TI-Nspire calculator is using the program Core Math Tools. You can download it from the website <u>http://www.nctm.org/coremathtools/</u> by clicking on "downloadable suite".

Once Core Math Tools is downloaded, open it and start a Simulation under Statistics & Probability.

- 1. From the column on the left, click the six-sided dice twice so we have two dice to work with.
- 2. Next highlight both die (on a PC this is done holding the "Control" key, while on a Mac you hold the "Command" key.)
- 3. Then, from the Build tab, click "Add" to simulate rolling two six-sided dice and adding the two together to show the sum.
- 4. We then want to see the number of times we roll the sum of an even number or a sum of 6. To simulate this, highlight the simulation, and click "Count # of" from the "Build" tab. A box will pop up with numbers 2-12. To count the number of even sums, you need to click on all the even numbers. First double click on the 2. Then, if you have a Mac, use "Command" while choosing "4". Continue doing that until you have highlighted all of the even numbers. On a PC, do the same thing but use "Control" instead of "Command".

- 5. Then, to find the number of times we roll the sum of an odd number or a sum of 6, we repeat the process above (begin with step #1) selecting 3, 5, 6, 7, 9, 11.
- 6. Finally, conduct 1000 Runs via the menu at the top. The simulation will show a "1" if the sum was even or 6 (odd or 6) and a "0" if not.
- 7. To find the total count of sums for each simulation, select "Graph" and "Labels Bars" in the "View" drop-down menu.



Christine Franklin Is ASA's First K-12 Statistical

Ambassador

The ASA is pleased to announce Christine Franklin (the lead author of the GAISE Pre-K–12 Report (*www.amstat.org/ education/gaise*) and the new *SET Report* (*www.amstat.org/education/set*), as the inaugural ASA K–12 statistical ambassador. Franklin will provide leadership in the creation and presentation of professional



development materials for teacher educators and teachers. She will present at national conferences, conduct workshops, collaborate with ASA chapters to enhance their education initiatives, and assist in outreach to the STEM education community. For more information, please see http://magazine. amstat.org/blog/2016/04/01/k12ambassador16.

ASA Issues Statistical Education of Teachers (SET) Report

www.amstat.org/education/set

The American Statistical Association (ASA) has issued the Statistical Education of Teachers (SET) Report, which calls on mathematicians, statisticians, mathematics educators, and statistics educators to collaborate in preparing pre-K-12 teachers to teach intellectually demanding statistics courses in their classrooms. SET was commissioned to clarify the recommendations for teacher statistical preparation in the Conference Board of the Mathematical Sciences' Mathematical Education of Teachers II report (http://cbmsweb.org/MET2/ *met2.pdf*). The SET report uses the ASA Pre-K–12 GAISE framework as the structure for outlining the content and conceptual understanding teachers need to know in assisting their students develop statistical reasoning skills. The report facilitates the understanding of key topics such as what sets statistics apart as a discipline distinct from mathematics, the difference between statistical and mathematical reasoning, and the role of probability in statistical reasoning. SET is intended for everyone involved in the statistical education of teachers, both the initial preparation of prospective teachers and the professional development of practicing teachers.

New ASA Stats 101 Toolkit Available

Many teachers of introductory statistics courses—whether at the high-school, two-year, or four-year college or university level—are trained in mathematics and may not have training or experience with statistics. At the request of 2015 ASA President David Morganstein, a group of statistical educators led by Dick De Veaux has written a series of case studies designed to show statistics in action, rather than as a branch of mathematics. Each case starts with a real-world problem and leads the reader through the steps taken to explore the problem, highlighting the techniques used in introductory or AP Statistics classes. Sometimes, the analysis goes slightly past the methods taught in such an introductory course, but the analysis is meant to build on simpler techniques and provide examples of real analyses, typical of the kind a professional statistician might perform. Our hope is that these case studies can both provide context and motivation for the instructor so the methods in the introductory course come alive. They can be used as examples in class, or just as guides for what a statistical analysis might entail. This repository can be found at *community.amstat.org/stats101/home*.

Teaching the Next Generation of Students to 'Think with Data'

How do we prepare students to engage with data in their work? That is the focus of a special issue of *The American Statistician*. The open-access guest editorial can be found at *http://bit.ly/22HWTzH*.

Significance Opens Archives

Significance magazine has opened its 10-year archives for access by the public. The magazine's volumes 1 through 10 are available to read, free of charge, at *www.statslife. org.uk/significance/back-issues.* Further, all magazine content will be made freely available one year after its initial publication. Editor Brian Tarran believes open access will demonstrate the importance of statistics and the contributions it makes in all areas of life.



Royal Statistical Society and ASA members and subscribers will continue to enjoy exclusive access to the latest magazine content.

Meeting Within a Meeting (MWM) Statistics Workshop for Middle- and High-School Mathematics and Science Teachers

Chicago, Illinois – August 2–3

MWM will take place in conjunction with the Joint Statistical Meetings this summer in Chicago, Illinois. The workshop is meant to strengthen K–12 mathematics and science teachers' understanding of statistics and provide them with hands-on activities aligned with the Common Core State Standards that they can use in their own classrooms. The cost of the workshop is \$50. Online registration is available at *www.amstat.org/education/mwm*.

Beyond AP Statistics (BAPS) Workshop

Chicago, Illinois – August 3

The ASA/NCTM Joint Committee is pleased to sponsor a Beyond AP Statistics (BAPS) workshop at the annual Joint Statistical Meetings. Organized by Roxy Peck, the BAPS workshop is offered for experienced AP Statistics teachers and consists of enrichment material just beyond the basic AP syllabus. The cost of the workshop is \$50. Online registration is available at *www.amstat.org/education/baps*.

CHANCE Special Issues on Nurturing Statistical Thinking Before College Parts 1 & 2 Available

The two-part (Nurturing Statistical Thinking Before College) *CHANCE* special issues are available at *bit.ly/21TQonH* and *bit.ly/1ZLaosY*. Some articles are freely available, while all articles are available through ASA Member's Only. Not yet an ASA K–12 Teacher Member? You can start your free one-year trial online at *www.amstat. org/membership/K12teachers.* There are also free recorded webinars accompanying four



articles from these special issues at *www. amstat.org/education/k12webinars.*

Special issue articles include the following:

- The Relationships Between Statistics and Other Subjects in the K–12 Curriculum, Zalman Usiskin
- The Statistical Education of Teachers: Preparing Teachers to Teach Statistics, Anna Bargagliotti & Christine Franklin
- The AP Statistics Exam: An Insider's Guide to Its Distinctive Features, Daren Starnes
- The AP Statistics Curriculum: Past, Present, and Future, Benjamin Hedrick
- A Catalyst for Change in the High-School Math Curriculum, Andrew Zieffler & Michael D. Huberty
- Data Surfing, Kay Endriss
- Taking a Chance in the Classroom: Professional Development MOOCs for Teachers of Statistics in K–12, Hollylynne Lee & Dalene Stangl

- STATS4STEM: Data, Computing, and Assessment Resources for High-School Statistics Students, Eric J. Simoneau
- Collaboration in the Mathematical Sciences Community on Mathematical Modeling Across the Curriculum, Peter R. Turner, Rachel Levy, & Kathleen Fowler
- Statistics: From College to Pre-College, Robert W. (Bob) Hayden
- Statistics in K–12: Educators, Students, and Us, Mary J. Kwasny
- Consider a Career in Statistics, John Holcomb
- Modeling Statistical Thinking, Daniel Kaplan
- Can 'Dirty' Data Be Your Friend?, Shep Roey, Tom Krenzke, & Robert Perkins

Data-Driven Mathematics Modules Available Free Online

Data-Driven Mathematics is a series of modules funded by the National Science Foundation and written by statisticians and mathematics teachers. Intended to complement a modern mathematics curriculum in the secondary schools, the modules offer materials that integrate data analysis with topics typically taught in high-school mathematics courses and provide realistic, real-world data situations for developing mathematical knowledge. The copyrights have been transferred from the original publisher to the ASA. Scanned copies of these books are freely available to download at *www.amstat.org/education/ publications.cfm*.

2016 Project Competition

Introduce your 7–12 students to statistics through the annual project competition directed by the ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability. The competition offers opportunities for students to formulate questions and collect, analyze, and draw conclusions from data. Winners will be recognized with plaques, cash prizes, certificates, and calculators. Also, their names will be published in *Amstat News*. Projects are due on June 1. For more information, visit *www.amstat.org/education/ posterprojects*. Note: There is an updated rubric for the project competition and new additional guidance under the project competition rules link posted at *www.amstat.org/education/ posterprojects*.

Judges Sought for Statistics Project Competition

The ASA/NCTM Joint Committee on Curriculum in Statistics and Probability is seeking judges for the 2016 Statistics Project Competition (*www.amstat.org/education/posterprojects*). Judging takes place via email during the summer and requires about four hours of your time. If interested, please email Nathan Kidwell, assistant head judge, at *nathan.kidwell@gmail.com*.

Stats.org Resources

Stats.org is a collaboration between the ASA and Sense About Science USA that aims to provide guidance regarding statistical literacy to journalists and the public. It provides interesting examples and stories that can be used in the classroom.

April is Mathematics Awareness Month

The 2016 Mathematics Awareness Month (MAM) theme is "The Future of Prediction." For the 2016 MAM poster and classroom resources, please visit *www.mathaware.org*.

ASA Updating GAISE College Report

The GAISE (Guidelines for Assessment and Instruction in Statistics Education) committee released a draft report, which provides an update of the recommendations for teaching introductory statistics in the 2005 *GAISE College Report*. The updated *GAISE College Report* is expected to be completed in fall 2016. The GAISE Pre-K–12 Report, the *GAISE College Report*, and the new draft report are available at *www.amstat. org/education/gaise*.

Free Online Videos and Courses for Statistics Teachers

YouTube Videos on Descriptive Statistical Concepts www.youtube.com/user/profdstangl/playlists

Need assistance in teaching your students statistical thinking? Through funding from Duke University and the American Statistical Association, Dalene Stangl, Kate Allman, Mine Cetinkaya-Rundel, and a group of Duke students have created a set of 52 videos to help you understand and teach basic descriptive statistical concepts. The videos are organized into five units. Within each, there are videos covering core concepts, pedagogy, JMP software, and applets. Unit one covers data and explains the structure of the videos. Unit two covers one-variable descriptive statistics, transforming a variable, and the normal curve. Unit three covers description of relationships between two categorical variables (contingency tables) and between one categorical and one numeric (side-by-side boxplots). Unit four covers description of relationship between two numeric variables using correlation and regression. Unit five pulls all the concepts

together in review videos.

Videos from *Data to Insight* Freely Accessible on YouTube

Chris Wild's *Data to Insight: An Introduction to Data Analysis* is a free, online, hands-on introduction to statistical data analysis. The videos that make up most of its "teaching content" have been made accessible on YouTube. While *Data to Insight* prototypes a next-generation introductory statistics course, many of its videos are immediately useful for current high-school and lower-level university statistics courses. The videos are indexed at *bit.ly/11BY4CZ* with an outline of their content and the course-design philosophy (see also the YouTube channel Wild About Statistics).

For those who actually want to learn about statistics, "materials" are only one part of the story. As we know, what learners do (activities) is more important than what they simply see. The course itself starts formally on October 19 and runs for eight weeks; however, it is self-paced and can be joined at any time until mid-December at *www.futurelearn.com/courses/data-to-insight*.

Against All Odds: Inside Statistics

www.learner.org/courses/againstallodds/index.html

Against All Odds is a free video series teaching introductory statistics concepts via real-life applications. This is an updated video series developed by Annenberg Learner (the producers of the original version in the 1980s) and contains videos, a glossary, teacher guides, and student guides.

Free Statistics Education Webinars

The ASA offers free webinars on K–12 statistics education topics at *www.amstat.org/education/webinars*.

This series was developed as part of the follow-up activities for the Meeting Within a Meeting Statistics Workshop (*www. amstat.org/education/mwm*). The Consortium for the Advancement of Undergraduate Statistics Education also offers free webinars on undergraduate statistics education topics at *www.causeweb.org.* Finally, the ASA/American Mathematical Association of Two-Year. Colleges (AMATYC) Joint Committee also offers free statistics webinars through AMATYC at *www. amatyc.org/?page=Webinars.*

Episode 17 of STATS+STORIES Is Available

Episode 17 ("A New Equation for Modern Journalism") of S+S is now available. Trevor Butterworth, director of Sense About Science USA and editor of Stats.org, joined the Stats+Stories regulars to discuss the skills now required for success in journalism. To listen, visit *www.statsandstories.net*. You can subscribe to the Stats+Stories podcast on iTunes.

Census at School Program Reaches More Than 40,000 Students

The ASA's U.S. Census at School program (*www. amstat.org/ censusatschool*) is a free, international classroom project that engages students in grades 4–12 in statistical problem solving. The students complete an online survey, analyze their class census results, and compare their class with random samples of students in the United States and other participating countries. The project began in the United Kingdom in 2000 and now includes Australia, Canada, New Zealand, South Africa, Ireland, South Korea, and Japan. The ASA is seeking champions to further expand the U.S. Census at School program nationally. For more information about how you can get involved, see the article at *bit.ly/1Uxq3wM* or email Rebecca Nichols at *rebecca@amstat.org*.

Explore Census at School Data with TuvaLabs

TuvaLabs provides free, real data sets, lessons, and visualization tools to enable teachers to teach statistics and quantitative reasoning in the context of real-world issues and topics. The ASA has provided TuvaLabs with a clean Census at School data set with 500 cases and 20 attributes that is freely available for students and teachers to explore online with their visualization tool and Census at School–adapted lesson plans. Start exploring Census at School data with TuvaLabs at *bit.ly/1LUGe4c*. Other TuvaLabs data sets and lessons are available at *www.tuvalabs. com*.

Online Community for ASA K–12 Teacher Members

A new online community for ASA K–12 Teacher Members will allow participation in online discussions and sharing resources with other members. More information is available at *http:// bit.ly/26tZvjS*. Not yet an ASA K–12 Teacher Member? You can start your free one-year trial at *www.amstat.org/membership/ K12teachers*.

World of Statistics Website and Resources

The free international statistics education resources created during the 2013 International Year of Statistics are available and ongoing through the World of Statistics website. Teachers everywhere can access a wealth of statistics instruction tools and resources from around the world at *www.worldofstatistics.org*.

PROJECT-SET

Project-SET is an NSF-funded project to develop curricular materials that enhance the ability of high-school teachers to foster students' statistical learning regarding sampling variability and regression. All materials are geared toward helping highschool teachers implement the Common Core State Standards for statistics and are closely aligned with the learning goals outlined in the *Guidelines for Assessment and Instruction in Statistics Education: A Pre-K–12 Curriculum Framework (GAISE).* For more information, visit *project-set.com*.

LOCUS Assessment Resources

LOCUS (locus.statisticseducation.org) is an NSF-funded project focused on developing assessments of statistical understanding across levels of development as identified in the Guidelines for Assessment and Instruction in Statistics Education (GAISE). The intent of these assessments is to provide teachers, educational leaders, assessment specialists, and researchers with a valid and reliable assessment of conceptual understanding in statistics consistent with the Common Core State Standards. The LOCUS website offers online assessment tools to measure statistical understanding. Create an account and manage test requests to receive immediate results from these automatically scored assessments. This is a great way for students to practice taking assessments online. Additional professional development resources are available with questions that are accompanied by high-quality commentaries written alongside student samples.

UPCOMING CONFERENCES

Electronic Conference on Teaching Statistics (eCOTS)

www.causeweb.org/cause/ecots/ecots16 May 16–20, 2016 (virtual conference)

Joint Statistical Meetings (JSM) *www.amstat.org/meetings/jsm*

July 30-August 4, 2016, Chicago, Illinois

Meeting Within a Meeting Statistics Workshop for Math and Science Teachers www.amstat.org/education/mwm August 2–3, 2016, Chicago, Illinois (JSM 2016)

Beyond AP Statistics (BAPS) Workshop www.amstat.org/education/baps August 3, 2016, Chicago, Illinois (JSM 2016)

U.S. News Rates Statistician as Best Business Job

The latest *U.S. News and World Report* Best Jobs Rankings named statistician as the best business job. In the overall 100 best job listings, statistician is #17 and, in best STEM job, #3. The statistician write-up quotes ASA member Devan Mehrotra, executive director of the biostatistics department at Merck Research Laboratories: "I absolutely fell in love with statistics. Any real-world problem almost always is going to require some data to be analyzed and interpreted, generating value-added solutions by using statistics. … There's a tremendous history … and now more exciting opportunities. It has never been a better time to be a statistician." *money.usnews.com/careers/bestjobs/statistician*

The ASA Releases Statement on Statistical Significance and *P*-Values

The American Statistical Association (ASA) has released a statement on statistical significance and *p*-values with six principles underlying the proper use and interpretation of the *p*-value (*bit.ly/21VdpMm*). The ASA releases this guidance on *p*-values to improve the conduct and interpretation of quantitative science and inform the growing emphasis on reproducibility of science research. The statement also notes that the increased quantification of scientific research and a proliferation of large, complex data sets has expanded the scope for statistics and the importance of appropriately chosen techniques, properly conducted analyses, and correct interpretation.

Common Core Resources

The Joint Committee of the American Statistical Association (ASA) and National Council of Teachers of Mathematics (NCTM) is committed to supporting the work of K–12 teachers. Carol Joyce Blumberg with assistance from Roxy Peck and Adam Molnar and oversight by the ASA/NCTM Joint Committee created a new resource that provides sources of lesson plans and learning activities to support the teaching of statistics and probability as covered by the Common Core State Standards. The document is divided into the following sections:

- American Statistical Association (ASA) Resources
- National Council of Teachers of Mathematics (NCTM) Resources
- Resources Developed by ASA Committees and Published Elsewhere
- Other Sources of Lesson Plans
- · Lesson Plans and Other Resources in Spanish
- Applets

- Videos
- Sources of Data
- Statistics and the Media
- Simulation Tools
- Random Number Generators and Random Samplers
- Technology (Including Statistical Calculators, Software Packages, & More)
- Resources for Teacher Preparation (Both Pre-Service and In-Service)
- Assessment Resources

The direct link for the document is *www.amstat.org/ education/pdfs/CommonCoreResources.pdf*.

Mobilize: Engaging Secondary Schools in Data Science

Robert Gould, Lead Principal Investigator

Statistics teachers know that requiring students to collect data is an effective way of engaging them in data analysis. Mobilize is an NSF-funded project that has developed a technology suite that allows students to engage in participatory sensing (PS) campaigns. In a PS campaign, students use their mobile devices to create a community that not only gathers data, but shares both the data and the analysis. The software includes an interactive "dashboard" that allows students to explore multivariate relations among variables that are numerical, categorical, location, data-and-time, text, and image.

The suite is used in a Mobilize-developed curriculum, the yearlong course Introduction to Data Science (IDS). IDS is a "C" approved mathematics course in the University of California A-G requirements, which means that successful completion of IDS validates Algebra II and provides an alternative admissions pathway to the University of California and California State University systems. To date, IDS has been taught in more than 35 classes and is still expanding.

The Mobilize project is a partnership between UCLA's Department of Statistics and Graduate School of Education and Information Studies and the Los Angeles Unified School District, the nation's second-largest school district. Mobilize is interested in identifying additional schools or districts that might be interested in adopting the IDS course or using the Mobilize technology suite. To learn more, contact LeeAnn Trusela, Mobilize project director, at *support@mobilizingcs.org* or visit the Mobilize website at *www.mobilizingcs.org*.

THIS IS STATISTICS

HELP US RECRUIT THE **NEXT GENERATION** OF STATISTICIANS

The field of statistics is growing fast. Jobs are plentiful, opportunities are exciting, and salaries are high. So what's keeping more kids from entering the field?

Many just don't know about statistics. But the ASA is working to change that, and here's how you can help:

- Send your students to *www.ThislsStatistics.org* and use its resources in your classroom. It's all about the profession of statistics.
- Download a handout for your students about careers in statistics at www.ThisIsStatistics.org/educators.

If you're on social media, connect with us at *www.Facebook.com/ThisIsStats* and *www.Twitter.com/ThisIsStats*. Encourage your students to connect with us, as well. The site features include:

- Videos of young statisticians passionate about their work
- A myth-busting quiz about statistics
- Photos of cool careers in statistics, like a NASA biostatistician and a wildlife statistician
- Colorful graphics displaying salary and job growth data
- A blog about jobs in statistics and data science
- An interactive map of places that employ statisticians in the U.S.



Lesson Plans Available on Statistics Education Web for K-12 Teachers

Statistics Education Web (STEW) is an online resource for peer-reviewed lesson plans for K–12 teachers. The lesson plans identify both the statistical concepts being developed and the age range appropriate for their use. The statistical concepts follow the recommendations of the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework, Common Core State Standards for Mathematics*, and *NCTM Principles and Standards for School Mathematics*. The website resource is organized around the four elements in the GAISE framework: formulate a statistical question, design and implement a plan to collect data, analyze the data by measures and graphs, and interpret the data in the context of the original question. Teachers can navigate the site by grade level and statistical topic. Lessons follow Common Core standards, GAISE recommendations, and NCTM Principles and Standards for School Mathematics.

Lesson Plans Wanted for Statistics Education Web

The editor of STEW is accepting submissions of lesson plans for an online bank of peer-reviewed lesson plans for K–12 teachers of mathematics and science. Lessons showcase the use of statistical methods and ideas in science and mathematics based on the framework and levels in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE)* and *Common Core State Standards*. Consider submitting several of your favorite lesson plans according to the STEW template to *steweditor@amstat.org*.

For more information, visit www.amstat.org/education/stew.

CENSUS at VSCHOOL

FREE international classroom project to engage students in statistical problem solving

Teach statistical concepts, statistical problem solving, measurement, graphing, and data analysis using your students' own data and data from their peers in the United States and other countries.

Complete a brief online survey (classroom census)

- 13 questions common to international students, plus additional U.S. questions
- 15–20-minute computer session

Analyze your class results

Use teacher password to gain immediate access to class data.

Formulate questions of interest that can be answered with Census at School data.

Collect/select appropriate data

Analyze the data—including appropriate graphs and numerical summaries for the corresponding variables of interest

Interpret the results and make appropriate conclusions in context relating to the original questions.

Compare your class census with samples from the United States and other countries

Download a random sample of Census at School data from United States students.

Download a random sample of Census at School data from international students (Australia, Canada, New Zealand, South Africa, and the United Kingdom).

International lesson plans are available, along with instructional webinars and other free resources.

www.amstat.org/censusatschool

For more information about how you can get involved, email Rebecca Nichols at *rebecca@amstat.org*.

Bridging the Gap Between Common Core State Standards and Teaching Statistics



Twenty data analysis and probability investigations for K-8 classrooms based on the four-step statistical process as defined by the Guidelines for Assessment and Instruction in Statistics Education (GAISE)

www.amstat.org/education/btg

Making Sense of Statistical Studies consists of student and teacher modules containing 15 handson investigations that provide students with valuable experience in designing and analyzing statistical studies. It is written for an upper middle-school or high-school audience having some background in exploratory data analysis and basic probability.

51

5/2

el I understand

everything.

9.000

www.amstat.org/education/msss

ST

a statistics book that makes sense!

4