Probability, Games and Sentiment Analysis
Nature Sunday Academy 2018-2019

Project Description:
Today we will cover the definition of *probability* and understand the nuances of this concept by repetitively performing simple experiments which have equally likely outcomes. We will also perform an experiment where the probability is skewed towards a particular outcome. Later, we will play two games: 1) *find the box with candy bag* – loosely based on Monty Hall show telecasted during 70’s; and 2) American roulette (straight-up bets and/ or black/ red bet) with fake currency. After the *find the box with candy bag* is over, the simple concept of probability will be used to educate students how they can improve their chances of winning a candy bag. After the students play American roulette, they will be taught how casinos make money and the common mistakes done by people while placing their bets.

In the second part of today’s activities, students will learn how online reviews can be misleading (numeric ratings versus actual reviews). They will learn about sentiment analysis that is popularly used by online retailers to better understand the feelings of their customers who previously purchased products from their websites. The students will be trained to draw word clouds of a product and a local restaurant that are more descriptive of the product/ business when coupled with numeric ratings. Overall, the students will learn to use the concept of probability in day-to-day life and will start viewing reviews available online with a different perspective after this workshop.

Project Objectives:
Objective 1: Introduce probability and use this concept to understand Monty Hall game and Roulette game.

Objective 2: Introduce sentiment analysis and use this concept to understand online reviews.

Learning Objectives:
Outcome 1: The participants will be able to use probability concepts in games and day-to-day life.

Outcome 2: The participants will be able to make better decisions based on the information they get from online resources.

Session Organization:

11:00-11:30am Cultural connection/ brief introduction
11:30-11:40am Introductions, definition of probability
11:40-11:55am Activity 1 (flipping coin 100 times and record # of heads/ tails, collect data and analysis)
11:55am-12:15pm Activity 2 (rolling dice 100 times and record # of 1’s, 2’s, 3’s, 4’s, 5’s and 6’s, collect data and analysis)
12:15-12:30pm Activity 3 (rolling biased dice 100 times and record # of 1’s, 2’s, 3’s, 4’s, 5’s and 6’s, collect data and analysis)
12:30-1:15pm Lunch
1:15-1:50pm Activity 4: Find the candy bag game (finding a box with candy box & explanation)
1:50-2:30pm Activity 5: roulette game (straight-up bet/ black/ red bets with fake currency & explanation)

2:30-2:40pm Sentiment analysis (introduction and procedure)

2:40-3:05pm Activity 6: review of online reviews

3:05-3:20pm Activity 7: build word clouds from reviews

3:20-3:30pm Conclude the workshop and provide printouts of word clouds prepared by students

**ND State Science Standards:**

8.2.2. Use evidence to generate descriptions, explanations, predictions, and models

9-10.2.1. Explain how scientific investigations can result in new ideas

9-10.2.3. Identify questions and concepts that guide scientific investigations

9-10.2.6. Design and conduct a guided investigation

9-10.2.7. Maintain clear and accurate records of scientific investigations

9-10.2.8. Analyze data found in tables, charts, and graphs to formulate conclusions

**Materials and Equipment:**

Print all the 6 worksheets and staple them together and provide them to all students.

Activity 1: quarters (number = students/ 2) and worksheet-1 (number = students/ 2)

Activity 2: dice (number = students/2) and worksheet-2 (number = students/2)

Activity 3: biased dice (number = students/2) and worksheet-3 (number = students/2)

Activity 4: three gift boxes; candy bags (number = # of players); worksheet-4 (number-1)

Activity 5: toy roulette, $1 bills fake currency (number = # of players × betting chances), worksheet-5

Activity 6: printouts of online reviews (number = # of 3 member groups)

Activity 7: desktops/ laptops/ tablets access to internet (number = # of 3 member groups)

**Activities:**

*Activity 1: flipping coin 100 times*

Step 1: ask students to find a working partner

Step 2: every pair receives one quarter and one worksheet-1

Step 3: one student flips the coin 100 times and the other student records the outcome each time

Step 4: each student pair will report the number of heads and tails along with probability for heads and tails.

Step 5: the moderator will add up the heads and tails from all the student pairs and will explain how the probability of heads/tails will converge to 0.5 when the data from all the students is gathered
**Activity 2: rolling dice 100 times**

Step 1: ask students to find a different working partner

Step 2: every pair receives one dice and one worksheet-2

Step 3: one student rolls the dice 100 times and the other student records the outcome each time

Step 4: each student pair will report the number of 1’s, 2’s, 3’s, 4’s, 5’s and 6’s along with probability for getting each number.

Step 5: the moderator will add up the number of 1’s, 2’s, 3’s, 4’s, 5’s and 6’s from all the student pairs and will explain how the probability of each number will converge to 1/6 when data from all the groups is collected and analyzed

**Activity 3: rolling biased dice 100 times (do not reveal that this is a biased dice)**

Step 1: ask students to find a different working partner

Step 2: every pair receives one biased dice and one worksheet-3

Step 3: one student rolls the biased dice 100 times and the other student records the outcome each time

Step 4: each student pair will report the number of 1’s, 2’s, 3’s, 4’s, 5’s and 6’s along with probability for getting each number.

Step 5: the moderator will add up the number of 1’s, 2’s, 3’s, 4’s, 5’s and 6’s from all the student pairs and will explain how the probability of each number will **NOT** converge to 1/6 in the case of biased dice.

**Activity 4: find candy bag game**

Step 1: choose a player

Step 2: place a candy bag in one of the three similar but numberered gift boxes and remember the number of this box. This should be done without the knowledge of the player. Other participants may know about this but should not help the player.

Step 3: the player chooses one of the three boxes.

Step 4: instead of revealing the contents of the box chosen by the player, the moderator reveals a box with no candy bag.

Step 5: Now, the moderator gives the player another chance. The player can either switch his/her option or stay with original choice. The moderator will record whether the player decides to switch or not in worksheet – 4.

Step 6: finally, the moderator reveals the box chosen by the player and if the box has a candy bag, the payer wins it.

Step 7: The moderator will ask the students play this game themselves using 3 cups and chocolate gold coins. Instructions provided on slide-8.
Step 8: the moderator will explain the participants how flipping will double their winning chances. Teach the logic for the second row in the table provided in slide 9 and ask the students help you to fill rest of the table. Using the details from the table find the probabilities of winning with and without switching. See handout on MH problem for solution.

**Activity 5: Roulette game**

Step 1: choose a player

Step 2: the player gets 5 $1 bills and will have 5 chances to bet on a number or a color.

Step 3: The losses and gains by the players will be recorded in worksheet – 5.

Step 4: After all the players are done playing, the moderator will explain the students how casinos make money. The moderator will also explain about gambler’s fallacy and reverse gambler’s fallacy.

More information about gamblers fallacy and reverse gamblers fallacy can be found here: [https://en.wikipedia.org/wiki/Gambler%27s_fallacy](https://en.wikipedia.org/wiki/Gambler%27s_fallacy)

**Activity 6: Analysis of reviews**

Step 1: The students will be divided in to several 3 membered groups.

Step 2: Each group will be given mixed online reviews of Iphone-7 and a local restaurant.

Step 3: The students will be asked to recognize the keywords used by reviewers to describe the product or service (find adjectives/ phrases that describe the product/ service).

Step 4: The students will be asked to look for trends (repeated phrases/ adjectives/ opinions)

**Activity 7: Building word clouds**

Step 1: open [www.wordclouds.com](http://www.wordclouds.com)

Step 2: enter keywords and the number of times they were repeated

Step 3: generate a word cloud of the two products and the business

**Wrap-up and discussion**

1) When an experiment is run finite times (less number of times), it may slightly deviate from theoretical probability. Ideally, an experiment has to be run infinite times (very high number of times) to hit theoretical probability.

2) Beware of biased coins, dices, roulettes etc.

3) The probability of winning a candy bag will double if one switches their choice after a box without candy is revealed.

4) By constantly placing bets on roulette for too many times, a player is bound to lose money.

5) Sentiment analysis can be used to better describe a product/ business.