

Teaching Statistics with Coins and Playing Cards – Going Beyond Probabilities

Authors

- Chris Malone (cmalone@winona.edu), Dept. of Mathematics and Statistics, Winona State University
- Tisha Hooks (thooks@winona.edu), Dept. of Mathematics and Statistics, Winona State University
- April Kerby (akerby@winona.edu), Dept. of Mathematics and Statistics, Winona State University

Note: Electronic version of this document provided at <http://course1.winona.edu/cmalone/research>

Abstract:

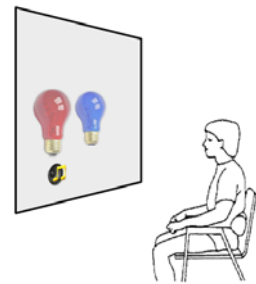
Traditionally, items such as coins and playing cards have been used to teach students about probability; however, these tools are typically not used in the introduction of concepts related to statistical inference. Take for example, the concept of a sampling distribution which is often introduced in a theoretical framework based on the normal distribution. Cobb (2007) argues that such an introduction is unnecessarily complicated. As a result, several authors (e.g. Tintle (2011), Rossman, Chance, Holcomb (2010)) have advocated for introducing the concepts of statistical inference through hands-on activities and simulations. In this presentation, we will share with you activities that we use in the classroom which use coins and playing cards to help introduce some of the more difficult concepts related to statistical inference.

Activity: Evaluating a Claim of Hearing Loss

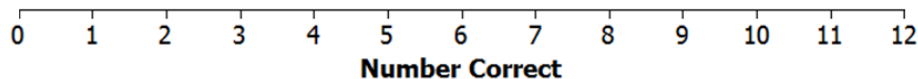
Consider the case study presented in an article by Pankratz, Fausti, and Peed titled “A Forced-Choice Technique to Evaluate Deafness in the Hysterical or Malingering Patient.” *Source: Journal of Consulting and Clinical Psychology, 1975, Vol. 43, pg. 421-422.* The following is an excerpt from the article:

The patient was a 27-year-old male with a history of multiple hospitalizations for idiopathic convulsive disorder, functional disabilities, accidents, and personality problems. His hospital records indicated that he was manipulative, exaggerated his symptoms to his advantage, and that he was a generally disruptive patient. He made repeated attempts to obtain compensation for his disabilities. During his present hospitalization he complained of bilateral hearing loss, left-sided weakness, left-sided numbness, intermittent speech difficulty, and memory deficit. There were few consistent or objective findings for these complaints. All of his symptoms disappeared quickly with the exception of the alleged hearing loss.

To assess his alleged hearing loss, testing was conducted through earphones with the subject seated in a sound-treated audiologic testing chamber. Visual stimuli utilized during the investigation were produced by a red and a blue light bulb, which were mounted behind a one-way mirror so that the subject could see the bulbs only when they were illuminated by the examiner. The subject was presented several trials on each of which the red and then the blue light were turned on consecutively for 2 seconds each. On each trial, a 1,000-Hz tone was randomly paired with the illumination of either the blue or red light bulb, and the subject was instructed to indicate with which light bulb the tone was paired.



Suppose that the subject was asked to do this experiment a total of 12 times. The possible outcomes for the number of correct answers will range from 0 to 12.



Questions:

1. What is the expected number of correct answers if the subject actually suffers from hearing loss and is therefore guessing on each trial?
2. Are there other values you would anticipate observing for the number of correct answers if the subject is guessing on each trial? What are these values?
3. What values would you anticipate observing if the subject were intentionally giving the wrong answer to make it look as though he couldn't hear?

Statistics can be used to determine whether or not a subject is intentionally giving the wrong answers. To investigate this situation, we can *simulate* the possible outcomes that a hearing impaired person would give for 12 trials of this experiment. For each replication of this experiment, we will keep track of how many times the subject was **correct**. Once we've repeated this process several times, we'll have a pretty good sense for what outcomes would be very surprising, or somewhat surprising, or not so surprising if the subject is truly hearing impaired.

Questions:

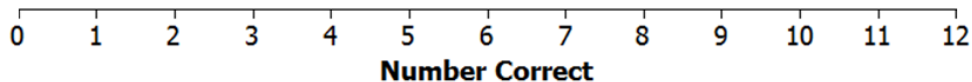
4. How can we replicate a hearing impaired individual using this applet?

5. Next, carry out 12 trials that simulate the responses of a hearing impaired individual and record your results below.

Trial	Choice		Correct?	
1	Red	Blue	Yes	No
2	Red	Blue	Yes	No
3	Red	Blue	Yes	No
4	Red	Blue	Yes	No
5	Red	Blue	Yes	No
6	Red	Blue	Yes	No
7	Red	Blue	Yes	No
8	Red	Blue	Yes	No
9	Red	Blue	Yes	No
10	Red	Blue	Yes	No
11	Red	Blue	Yes	No
12	Red	Blue	Yes	No

How many correct answers did you have? _____

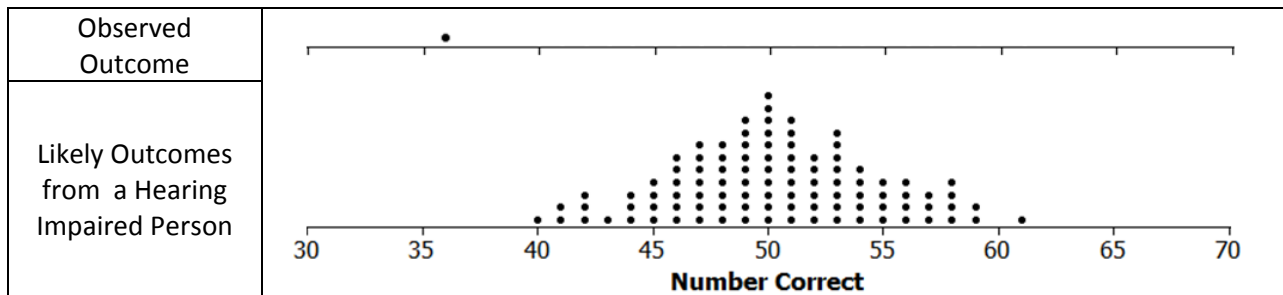
Next, collect the simulation results from everybody in the class. Make a dot for your outcome and that of all others in your class on the following number line.



Questions:

6. Given the simulation results on the above dotplot, what would you think about the subject's claim that he suffered hearing loss if he answered ...
- a. 5 correctly?
 - b. 0 or 1 correctly?
 - c. 2 or 3 correctly?

In the actual study, the subject was asked to complete 100 trials (instead of 12 trials as used above). The graphic below was obtained using a computer to simulate the possible outcomes of a hearing impaired person (i.e. guessing). Each time we simulated the experiment we counted and recorded the number of correct answers. This process was repeated several times and the results are shown below.





Questions:

7. The subject gave the correct answer in 36 of the 100 trials. What do you think about the subject's claim that he suffers from hearing loss?

Activity: Making Decisions in Criminal Investigations

Suppose that in the course of a criminal investigation, detectives develop a series of 20 true-false questions about the crime scene, for example, and then embed these questions in the interrogation of a suspect. The suspect is not allowed to say, "I don't know;" instead, the detectives force the suspect to answer "true" or "false" for each of the 20 questions.

A coin can be used to simulate how a suspect who has no knowledge of the crime and is truly guessing might answer these 20 true-false questions.

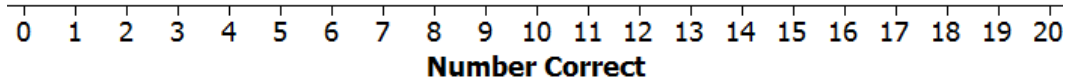
	Suspect guesses <u>correctly</u>
	Suspect guesses <u>incorrectly</u>

Questions:

Coin Model	Guessing Model
1a) If you toss a coin 20 times, how many coins would you expect to land on heads?	1b) If a suspect has no knowledge of the crime and is truly guessing on 20 true-false questions, how many would you expect them to answer correctly?
2a) A classmate tosses a coin 20 times and gets 9 heads. They claim their coin is biased towards tails. What is wrong with their reasoning?	2b) A suspect answers 9 out of the 20 true-false questions correctly. The investigators claim that since this was less than the expected number of correct answers, the suspect must be answering incorrectly on purpose. What is wrong with their reasoning?
3a) A classmate tosses a coin 20 times and gets 0 heads. They claim their coin is biased towards tails. Do you agree with their reasoning?	3b) A suspect answers 0 out of the 20 true-false questions correctly. The investigators claim that the suspect must be answering incorrectly on purpose. Do you agree with their reasoning?

Questions:

4. In your opinion, at what point would you become convinced that the suspect is giving wrong answers intentionally?



5. Ask some of your neighbors at what point they would become convinced the suspect is giving wrong answers intentionally.



Neighbor 1: _____ Neighbor 2: _____

Neighbor 3: _____ Neighbor 4: _____

6. What potential issues arise when different people have different opinions regarding when they become convinced the suspect is giving wrong answers intentionally?

Simulation Setup

To simulate the situation where the suspect has no knowledge of the crime and is guessing, we will flip a fair coin 20 times, one flip for each true-false question the suspect answers.

	Suspect guesses <u>correctly</u>
	Suspect guesses <u>incorrectly</u>

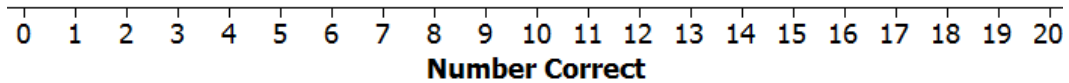
Record the outcome from each coin toss in the table below.

Coin Toss	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Outcome																					

Count the total number of heads (i.e. correct answers) from your 20 coin flips.

Number of correct answers: _____

Make a dot for your outcome and that of all others in your class on the following number line.



Questions:

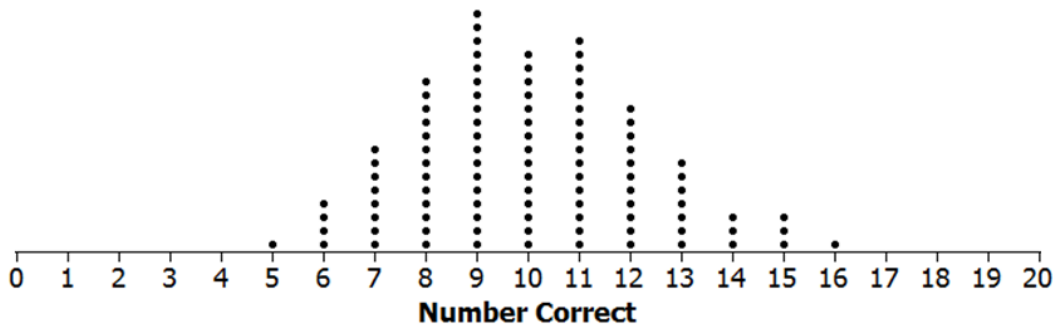
7. What does each dot on the above dotplot represent?

8. Based on these simulated results, if a suspect answered only 8 questions correctly, would you believe they were intentionally giving wrong answers?

9. If a suspect answered only 3 questions correctly, would you believe they were intentionally giving the wrong answers?

10. After seeing these simulated results, at what point would you become convinced that the suspect is giving incorrect answers on purpose?

A computer can be used to mimic the process of flipping a coin 20 times, over and over again. This will improve our ability to make good sound decisions regarding a suspect's possible knowledge of a crime.



Questions:

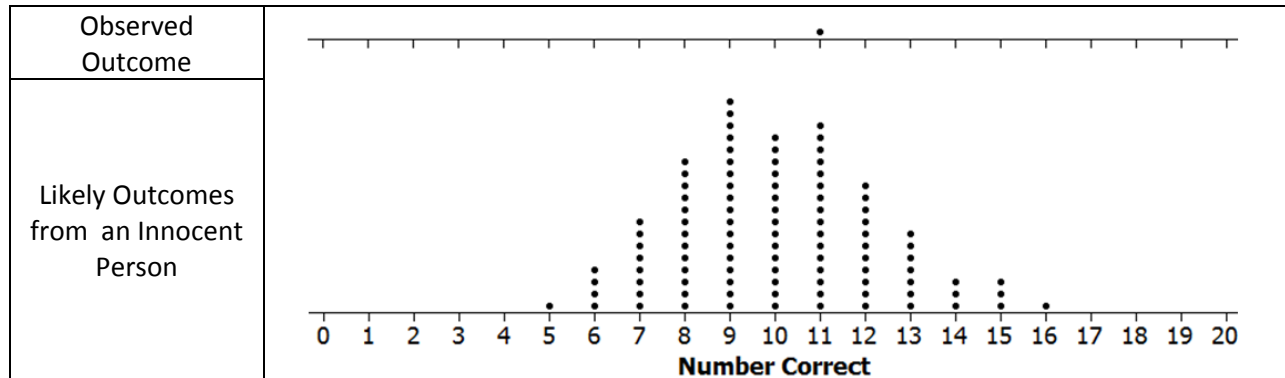
The following scenarios are based on actual case studies discussed in a forensic science publication.

Source: Harold V. Hall and Jane Thompson. "Explicit Alternative Testing: Applications of the Binominal Probability Distribution to Clinical-Forensic Evaluations." *The Forensic Examiner*, Spring 2007.

11. A suspect in a rape/murder case bragged about committing the crime and disposing of the victim's body. Police embedded 20 true-false questions concerning the victim's characteristics and the crime scene in an interrogation. The suspect answered 2 of the 20 questions correctly. Based on the above dotplot, would you believe he had knowledge of the victim and crime scene and was trying to hide it?

Observed Outcome	
Likely Outcomes from an Innocent Person	

12. During the early portion of the Iraq War, a lieutenant was suspected of reading and then stealing war plans with which he was entrusted to transport between headquarters. The division's lead intelligence officer who had read these plans helped devise 20 true-false questions that only someone who had read the plans would be able to answer with confidence. The lieutenant answered 11 of the 20 questions correctly. Based on the above dotplot, would you believe he was intentionally trying to hide knowledge of the war plans by giving incorrect answers on purpose?



A Common Standard for Making Decisions

The further we move towards the lower end of this distribution, the more evidence we have that the suspect is lying. When concern lies in the lower end of the distribution, statisticians would conclude we have enough evidence that the suspect is lying when the observed number of correct guesses falls in the bottom 5% of the distribution.



Activity: Vested Interest and Task Performance

This example is from Investigating Statistical Concepts, Applications, and Methods by Beth Chance and Allan Rossman. 2006. Thomson-Brooks/Cole.

A study published in the *Journal of Personality and Social Psychology* (Butler and Baumeister, 1998) investigated whether having an observer with a vested interest would decrease subjects' performance on a skill-based task; in other words, researchers suspected that those who were playing for both themselves and an observer would be more likely to "crack" under the pressure. Subjects were given time to practice playing a video game that required them to navigate an obstacle course as quickly as possible. They were then told to play the game one final time with an observer present, and subjects were randomly assigned to one of two groups. The Vested Interest group was told that the participant and observer would each win \$3 if the participant beat a certain threshold time; the No Vested Interest group was told that the participant only would win \$3 if the threshold were beaten. The threshold was chosen to be a time that each participant beat in 30% of their practice turns.

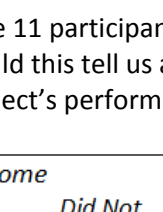




Investigating Possible Outcomes

Twelve participants were randomly assigned to each group in this study. A total of 11 participants beat the threshold. The data from this study can be organized as follows.


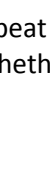

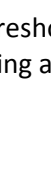


		<i>Outcome</i>		
		<i>Beat Threshold</i>	<i>Did Not Beat Threshold</i>	
<i>Vested Interest</i>				12
<i>No Vested Interest</i>				12
		 11	 13	

Questions:







1. What would the following outcome tell us about whether having an observer with a vested interest would decrease a subject's performance?

	Outcome		
	Beat Threshold	Did Not Beat Threshold	
Vested Interest			12
No Vested Interest			12
			

2. Consider the following outcome in which the 11 participants that beat the threshold are divided equally between the two groups. What would this tell us about whether having an observer with a vested interest would decrease a subject's performance?

	Outcome		
	Beat Threshold	Did Not Beat Threshold	
Vested Interest			12
No Vested Interest			12
			

3. The next table contains the actual outcomes from the study. Do these results convince you that having an observer with a vested interest decreases a subject's performance?

	Outcome		
	Beat Threshold	Did Not Beat Threshold	
Vested Interest			12
No Vested Interest			12
			

Comments:

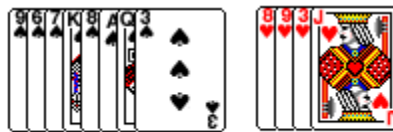
- Note that given both the group totals and the overall number of participants that beat the threshold, we really need only the number that beat the threshold in the Vested Interest group to fill in the rest of the table.
- As stated earlier, if having an observer with a vested interest present really has no impact on task performance, we expect our participants who beat the threshold to be split evenly between Vested Interest and No Vested Interest groups. However, note that *by pure chance alone*, we could end up with fewer participants who beat the threshold in the Vested Interest group just because of the way that random assignment to groups works out.
- The smaller the number of participants that beat the threshold in the Vested Interest group, the more we believe that being in the group with a Vested Interest decreases task performance. *How small does this value have to get in order for us to be convinced it is not happening just because of random chance?* We will answer this question by simulating this experiment over and over again, but in a situation where we know that being in the group with Vested Interest has no effect.

A Simulation Study with Playing Cards

To determine how we can expect the counts in each group to end up just by chance, we could let 11 red cards represent those participants that beat the threshold, and let 13 black cards represent those that did not beat the threshold.

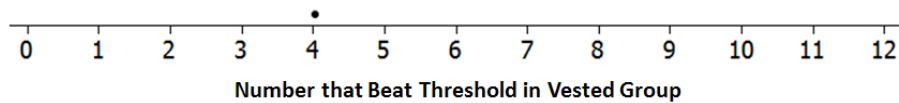


After shuffling the cards well, suppose we randomly deal out 12 to represent the Vested Interest group and the other 12 to represent the No Vested Interest group (note that this mimics the random assignment of participants to groups).

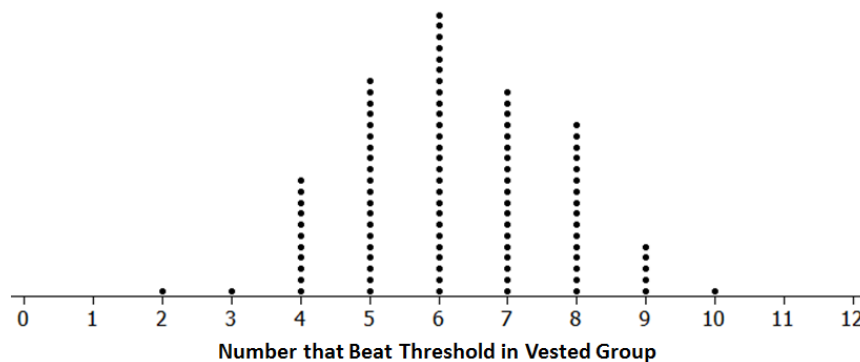


Pile for the Vested Interest group

Then, we count the number of red cards (which represent those that beat the threshold) in the Vested Interest group and record this value on the plot below.



Next, suppose that we repeated this randomization process 99 more times, each time counting the number that beat the threshold in the Vested Interest group and recording that value on the plot.



Questions:

- Recall that in the actual study only 3 of the 11 participants that beat the threshold were in the Vested Interest group. Does this convince you that having an observer with a vested interest present decreases a subject's performance? Explain.

Activity: Effect of Wording on Survey Questions

Question wording on a survey can sometimes impact the response. For example, consider two different wordings of a question commonly asked on the General Social Survey:

Version 1	As a country, do you think we are spending too little money on assistance to the poor?
Version 2	As a country, do you think we are spending too little money on welfare?

Even though the questions are worded differently, the meaning of the question is the same in both cases. To see if the alteration of words has an impact on how people respond to the question, we will carry out a study in which you are all subjects. Some of you were randomly assigned to Version 1 of the question, and others were randomly assigned to Version 2. Obtain the following summaries.

- Number who answered “assistance to poor” version: _____
- Number who answered “welfare” version: _____
- Number who answered Yes: _____
- Number who answered No: _____

Questions:

1. Put the above summaries in the following table. We suspect that the word “welfare” will produce more negative responses. Give an example of what the table would look like if this theory is correct.

Assuming the word “welfare” has a negative effect

Group	Do you think we are spending too little money?		Total
	Yes	No	
Assistance to the poor			
Welfare			
Total			

2. Give an example of what the table would look like if the wording of the question has no effect on how a person responds.

Assuming the word "Welfare" has no effect

Group	Do you think we are spending too little money?		Total
	Yes	No	
Assistance to the poor			
Welfare			
Total			

3. Give an example of what the table would look like if the wording of the question has a borderline effect on how a person responds.

Assuming the word "Welfare" has a borderline effect

Group	Do you think we are spending too little money?		Total
	Yes	No	
Assistance to the poor			
Welfare			
Total			

Setting up the Simulation Study

We will use playing cards to replicate this experiment under the assumption that a subject who answered “Yes” was equally likely to have come from either group.

- There are _____ subjects in our study that answered “Yes” when asked if we spend too little. Use red cards to represent these subjects.
- There are _____ subjects in our study that answered “No” when asked if we spend too little. Use black cards to represent these subjects.
- Shuffle the cards together well, and randomly deal out a pile of _____ cards to represent those assigned to the “Welfare” group. Then, record the number of red cards (representing those that answer “Yes”) that ended up in this pile by random chance. Record this number below.

Number of subjects that answered “Yes” in the “Welfare” group in your simulation: _____

Next, collect the simulation results from everybody in the class. Make a dot for your outcome and that of all others in your class on the following number line.

Number that answer Yes in Welfare group

Questions:

4. The table of counts obtained in our actual study will now be revealed. How many subjects answered “Yes” in the Welfare group?

Actual Counts

Group	Do you think we are spending too little money?		Total
	Yes	No	
Assistance to the poor			
Welfare			
Total			

5. Based on the results of this simulation study, does it appear these results were likely to have happened by random chance alone? Explain.

6. Would you say that the data provide convincing evidence that the word "Welfare" produces more negative responses? Explain.

Source: Tom Smith. *"That Which We Call Welfare By Any Other Name Would Smell Sweeter: An Analysis of the Impact of Question Wording on Response Patterns."* *Public Opinion Quarterly* 1987, Volume 51: 75 - 83.