Algebra 2

ELG HS.S.4: Understand and evaluate random processes underlying statistical experiments.

Vertical Progression:

7th Grade	**7.SP.A Investigate patterns of association in bivariate data** o **7.SP.A.1** Understand that statistics can be used to gain information about a population by examining a sample of the population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences. o **7.SP.A.2** Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions.
Algebra 2	**ELG.MA.HS.S.4 Understand and evaluate random processes underlying statistical experiments.** o **S-IC.1** Understand statistics as a process for making inferences about population parameters based on a random sample from that population. o **S-IC.2** Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation. *For example, a model says a spinning coin falls heads up with probability 0.5. Would a result of 5 tails in a row cause you to question the model?*

Students will demonstrate command of the ELG by:

- Defining populations, population parameter, random sample, and inference.
- Explaining why randomization is used to draw a sample that represents a population well.
- Recognizing that statistics involves drawing conclusions about a population based on the results obtained from a random sample of the population.
- Comparing theoretical and empirical results to evaluate the effectiveness of a treatment.
- Explaining how well and why a sample represents the variable of interest from a population.
- Demonstrating understanding of the different kinds of sampling methods.
- Designing simulations of random sampling, assigning digits in appropriate proportions for events, carrying out the simulation using random number generators and random number tables and explaining the outcomes in context of the population and the known proportions.

Vocabulary:

- Empirical results
- Inference
- Population
- Population parameter
- Randomization
- Random number generators
- Random number tables
- Random sample
- Sample
- Sampling methods
- Theoretical results
- Treatment

Sample Instructional/Assessment Tasks:

1) **Standard(s): S-IC.1**

Source: https://www.illustrativemathematics.org/content-standards/tasks/186

Item Prompt:

From a class containing 12 girls and 10 boys, three students are to be selected to serve on a school advisory panel. Here are four different methods of making the selection.

I. Select the first three names on the class roll.

II. Select the first three students who volunteer.

III. Place the names of the 22 students in a hat, mix them thoroughly, and select three names from the mix.

IV. Select the first three students who show up for class tomorrow.

Which is the best sampling method, among these four, if you want the school panel to represent a fair and representative view of the opinions of your class. Explain the weaknesses of the three you did not select as the best.

Correct Answer:

Choice III is the best solution in terms of fairness because each of the other methods does not give equal chance of selection to all possible groups of three students. Explanations as to why the others are unfair may include comments such as the following:

I. Names beginning with the same letter may belong to the same family or the same ethnic group.

II. Volunteers may have special interest in a particular issue on which they want to focus.

IV. Prompt students may be the more serious students and, perhaps, would be the more conscientious members of a panel, but they may not be typical of students in the class.

2) **Standard(s): S-IC.S**

Source: https://www.illustrativemathematics.org/content-standards/HSS/IC/A/2/tasks/1099

Item Prompt:

Many researchers have studied chimpanzees to learn about their problem solving skills. In 1978, researchers Premack and Woodruff published an article in Science magazine, reporting findings from a study on an adult chimpanzee named Sarah, who had been raised in captivity and had received extensive training using photos and symbols. In one experiment, Sarah was shown videotapes of eight different situations in which a human being was faced with a problem. After each videotape showing, Sarah was presented with two photographs, one of which depicted a possible solution to the problem. The researchers ensured that the order in which the photographs were presented was randomized (for example, the correct answer was not always presented first, etc.) and that the photographs had similar visuals (for example. similar colors, etc.) Of the eight problems, Sarah picked the photograph with the correct solution seven times. Could Sarah have been merely guessing and just lucky with her responses, or is there evidence that Sarah does better than just guessing?

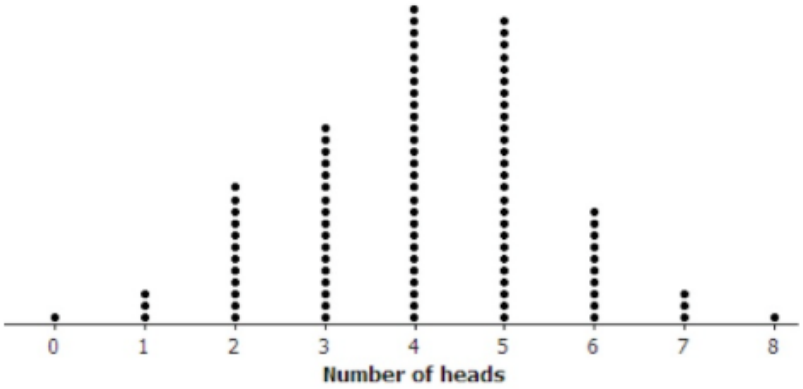
a. Give two *possible* explanations for why Sarah might have answered seven out of eight correctly.

b. If Sarah were just guessing, and was just likely to pick one photograph compared to the other, how many would you expect her to get right out of eight problems?

c. Give an example of how you could use basic classroom tools (coins, dice, calculators, cards, etc.) to simulate one trial of Sarah "just guessing" to pick a photograph for one problem.

d. A student, James, decides to use simulation to investigate whether the study data provide evidence that Sarah was doing better than just randomly guessing, and so James tosses a coin eight times, and obtains six heads. Explain why James should repeat the process of tossing the coin eight times and recording the number of heads, many times.

e. James repeats the process of eight coin-tosses 100 times, each time recording the number of heads on the eight coin-tosses. The following is a dotplot of his results.



Based on the above dotplot, what was the most common result for "number of heads" in eight coin-tosses? Why does this make sense?

f. Based on this dotplot, would you say that a score of 7 out of 8 would be unusual if Sarah has just been guessing? Why or why not?

g. Which of the following is a *possible* explanation for Sarah's performance?

i. Sarah had been just guessing and got lucky with her responses.

ii. Sarah does better than just guessing.

iii. Both (i) and (ii) are possible explanations.

h. Based on the simulation results, which of the following appears to be a *plausible* (likely) explanation for Sarah's performance?

i. Sarah had been just guessing and got lucky with her responses.

ii. Sarah does better than just guessing.

iii. Both (i) and (ii) are possible explanations.

i. Based on the results of this study, would it be reasonable to say that all chimpanzees do better than just guessing when faced with the kind of problems posed to Sarah? Explain why or why not.

Correct Answer:

a. Either Sarah was randomly guessing and got lucky *by chance*, or Sarah does better than randomly guessing.

b. If Sarah had been guessing, we would expect her to get 4 (half of eight) correct. (Instead she got seven right. The question is, *Is seven out of eight different enough from four out of eight to conclude that Sarah was doing better than just guessing?*)

c. Use a coin flip; heads for correct answer and tails for incorrect answer. Or, a die; even number for correct answer and odd for incorrect.

d. Because James needs to see what happens in the *long-run* on eight coin-tosses. What outcomes are more common? What outcomes are less common? How often does 7 heads in eight coin-tosses happen just by chance?

e. Most common outcome is 4 - which makes sense because that is what we expect would happen if Sarah was randomly picking a photograph.

f. 7 is unusual (surprising) because on the 100 tosses, an outcome as or more extreme as 7 happened by chance only 4 times.

g. (iii) Both (i) and (i) are possible explanations

h. (ii) Sarah does better than just guessing.

i. The question of interest is about Sarah getting the answer right, rather than about all chimpanzees. Note that Sarah's trials are not a random sample from the population of all possible chimp responses.