# Investigation 16

Too Many Peanuts?

*Investigating a Claim*

## Overview

This investigation introduces the concept of informal statistical inference. Using technology, students construct a sampling distribution of sample proportions to determine if an observed sample proportion would be considered unusual for a given population proportion. Students will be testing the claim that cans of mixed nuts contain approximately 50% peanuts. 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 statistical question, design and implement a plan to collect data, analyze the data by measures and graphs, and interpret the results in the context of the original question. This is a GAISE Level C activity.

*Note:* If any students have an allergy to peanuts, rather than open a can of nuts, use the data presented in the lesson.

## Instructional Plan

### Brief Overview

» Develop a statistical question about the proportion of peanuts in a can of mixed nuts that is claimed to contain approximately 50% peanuts.

» Prior to class, count the number of nuts in the can and consider that number to be the random sample size of mixed nuts taken from the population of all mixed nuts processed by the manufacturer. In this investigation, the number of mixed nuts is 258.

» Calculate the proportion of peanuts in the sample. In this investigation, the proportion of peanuts in the sample of 258 mixed nuts is 55% peanuts.

» Construct a sampling distribution of sample proportions of size (number of nuts in the can). In this investigation, 258 mixed nuts from a population with a proportion of 50% peanuts is used as an example.

» Based on the sampling distribution, find the probability of randomly obtaining a sample proportion of peanuts of at least 55% peanuts, assuming the population from which the sample is taken contains 50% peanuts.

Hand out Student Worksheet 16.1 Peanut Investigation.

Ask your students to read the scenario.

### Scenario

Did you ever buy a can of mixed nuts and it seemed all you got in the can was peanuts and you were hoping for a lot of cashews and almonds?

A 1964 *Consumer Reports* investigation of 124 cans of mixed nuts, representing 31 brands bought in 17 American cities, determined that most mixed nuts at that time were mostly peanuts, often 75%. As of 1993, the Food and

Drug Administration (FDA) has required a container of mixed nuts to contain at least four varieties of tree nuts or peanuts. Each kind of nut must be present not less than 2% and not more than 80% of the number of nuts.

A major manufacturer of cans of mixed nuts makes the claim that their 10.3 oz. cans containing a mixture of peanuts, almonds, cashews, pecans, and Brazil nuts have approximately 50% peanuts.

As part of a statistics project, an 11th grader purchased a 10.3 oz. can of mixed nuts and found 142 peanuts in the can that contained 258 mixed nuts or approximately 55% peanuts.

**Learning Goal**

Use the sampling distribution of sample proportions and informally decide if a single sample proportion is unusual.

**Mathematical Practices Through a Statistical Lens**

*MP3. Construct viable arguments and critique the reasoning of others.*

Statistically proficient students use appropriate data and statistical methods to draw conclusions about a statistical question. They reason inductively about data, making inferences that take into account the context from which the data arose. They justify their conclusions and communicate them to others.

**Materials**

Student worksheets are available at *www.statisticsteacher.org/statistics-teacher-publications/focus.*

» 10.3 oz. can of mixed nuts that the manufacturer claims contains approximately 50% peanuts

» Statistical software or application to generate a sampling distribution of sample proportions. Possible applications: Graphing calculator with ProbSim app or computer software like GeoGebra or StatKey

» Student Worksheet 16.1 Peanut Investigation

» Exit Ticket

» Optional: Student Worksheet 16.2 StatKey Directions

**Estimated Time**

One 50-minute class

**Pre-Knowledge**

Students should be able to find the mean and standard deviation of a distribution using technology.

Does this mean the manufacturer's claim of approximately 50% peanuts is not correct? Does this provide convincing evidence that cans of mixed nuts from this manufacturer contain more than 50% peanuts?

*Note:* If appropriate, open the can of mixed nuts and count the total number of nuts and the number of peanuts. Determine the proportion of peanuts in the can. Use these results to complete this investigation.

*Note:* If there are students with peanut allergies, then use the 55% result computed from a can containing 258 nuts. This investigation will use the 55% results from a can containing 258 nuts as the example.

### Formulate a Statistical Question

Start the discussion by explaining that we are going to assume the claim of approximately 50% of mixed nuts are peanuts is true. Another way to say this is the population proportion of peanuts in all the mixed nuts is approximately 50%. Explain that we are also going to assume the number of nuts in the can of mixed nuts represents a random sample from a population of all mixed nuts produced by this manufacturer.

Next, ask your students, "If the manufacturer's claim of 50% peanuts is true, how likely is it that we get a can of mixed nuts that contains 55% peanuts? Is this an unusual result? What is the probability we could get a sample containing 55% peanuts by chance from a population containing 50% peanuts?"

Ask your students to consider the statistical question: "Assuming the manufacturer's claim that a can of mixed nuts contains 50% peanuts is true or the population proportion is 0.5, is the proportion of peanuts of 0.55 found in a can (sample) of mixed nuts an unusual result?"

### Collect Appropriate Data

Ask your students to answer questions 1 to 3.

We are going to assume the population proportion of peanuts in all the mixed nuts processed by the manufacturer is 50% and the sample of 258 nuts (one can) was a random sample of all the mixed nuts produced by the manufacturer.

1. Assuming the claim that 50% of a can is peanuts, how many peanuts would you expect to be in a can of 258 nuts?

***Answer:*** *Fifty percent of 258 is 129.*

2. If a random sample of 258 mixed nuts yielded 134 peanuts, would you think it an unusual result? Why or why not?

***Answer:*** *This would not be considered unusual, since 134/258 is approximately 52%, which is close to 50%.*

3. If a random sample of 258 mixed nuts yielded 155 peanuts, would you think it an unusual result? Why or why not?

***Answer:*** *This is a very unusual result. 155/258 is about 60%, which is much higher than the claim of 50%.*

Ask your students to complete questions 4 and 5.

*Note:* The following results were constructed using the statistical computer application-StatKey (*www.lock5stat.com/StatKey*).

Steps for using StatKey are on Student Worksheet 16.2 StatKey Directions.

4. As directed by your teacher, use statistical software to construct a simulated sampling distribution of at least 200 sample proportions based on a sample size of 258—number of nuts in the can—and assuming a population proportion of 50%.

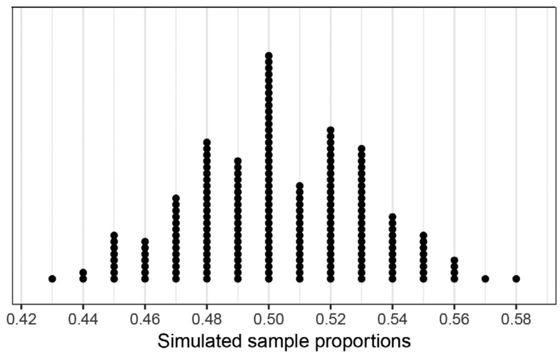


Figure 16.1: Sampling distribution of 200 sample proportions

***Sample answer:*** *Figure 16.1*

## Analyze the Data

Ask your students to answer questions 5 to 9.

5. What do you expect the mean of the simulated sampling distribution of sample proportions to equal?

***Answer:*** *Approximately 0.50*

6. Using statistical software, find the mean and standard deviation of the simulated sampling distribution.

***Sample answer:*** *Mean = 0.501, or 0.5; standard deviation = 0.030*

*Note:* The standard deviation of the sampling distribution is called the standard error of the sample proportion.

7. Describe the simulated sampling distribution of the sample proportions.

***Sample answer:*** *Mound shaped or approximately Normal with a mean of approximately 0.5 and a standard deviation of approximately 0.03. Most sample proportions are between 0.44 and 0.56.*

8. Count the sample proportions on the plot that are greater than or equal to the proportion of peanuts in the class can (0.55 in this example). How many sample proportions were greater than or equal to the class proportion of peanuts?

***Sample answer:*** *14 out of 200 (based on this example)*

9. Estimate the probability of the class getting a can of mixed nuts and obtaining a sample proportion of __% (in this example, 55%) peanuts or greater from a population with the population proportion equal to 0.50 peanuts.

***Sample answer:*** *14/200 or 7% chance (based on this example)*

## Interpret the Results in the Context of the Original Question

Ask the students to answer questions 10 to 12.

10. Do you think the proportion of peanuts in the class can of mixed nuts was an unusual result assuming the manufacturer's claim of 50% is correct?

***Answer:*** *Since an estimate for the probability of obtaining the class sample proportion is 7% (answer and interpretation will vary based on the class results), the sample proportion is not that unusual. There is "some" evidence that the number of peanuts is higher than the usual amount, but there is not enough evidence to say the company's claim of 50% peanuts is not true.*

11. What proportion of peanuts would you consider to be an unusual result? Based on the simulated sampling distribution, what is an estimate for the probability of obtaining that proportion or more by chance?

***Possible answers:*** *Answers will vary, but students may respond with a sample proportion of around 57% or higher. In this example, the probability is approximately 2/200, or 0.01.*

12. If you got such a can (high proportion of peanuts), would you have reason to believe the manufacturer's claim is not correct?

***Possible answer:*** *Even though it can happen, the probability is very low, and I would not believe the manufacturer claim of 50% peanuts.*

### Additional Ideas

» Use survey results from your class or classes and test the claim that 75% of teens use Snapchat.

» Use survey results from your class or classes and test the claim that 50% of teens use Twitter.

» Use survey results from your class or classes and test the claim that fewer than 30% of teens use Tumblr, Twitch, or Linkedln.

## Exit Ticket

The American Society for the Prevention of Cruelty to Animals (ASPCA) claims that approximately 35% of US households have at least one cat. Assuming the ASPCA's claim is correct, a sampling distribution of 100 sample proportions based on a sample size of 50 and population proportion of 0.35 is shown in Figure 16.2.

Mean of simulated sampling distribution = 0.35 and a standard deviation = 0.06

1. Describe the simulated sampling distribution of the sample proportion.

***Answer:*** *The distribution is mound shaped or approximately Normal with a mean of 0.35 and standard deviation of 0.06.*

2. A random sample of 50 US households found the proportion of households with at least one cat was 0.22. Mark the sample result of 0.22 on the dot plot. How many sample proportions are less than or equal to the sample result of 0.22?

***Answer:*** *Approximately 3 out of 100, as shown in Figure 16.3*

3. What is an estimate for the probability of obtaining a sample proportion of 0.22 or less from a population with 0.35 households with a cat?

# Exit Ticket Cont.

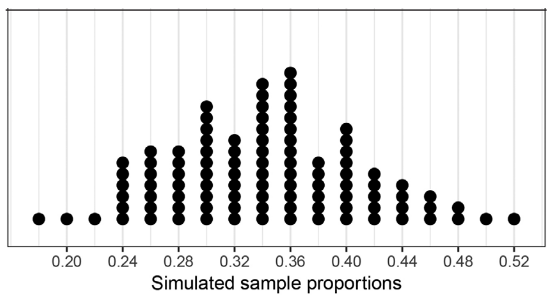


Figure 16.2: A sampling distribution of 100 sample proportions based on a sample size of 50 and population proportion of 0.35

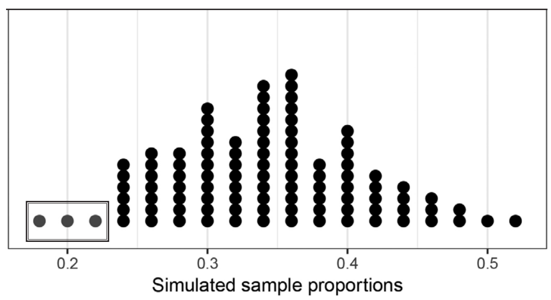


Figure 16.3: Sample result

***Answer:*** *Approximately 3/100 or 0.03*

4. Do you think the proportion of households with a cat (22%) was an unusual result, assuming the ASPCA's claim of 35% is correct? Explain your answer.

***Possible answer:*** *The probability of obtaining a sample result of 0.22 households with a cat from a population with a proportion of 0.35 households with a cat is about 3%. This is a low probability, so I think this is an unusual result.*

## Further Exploration and Extensions

1. Introduce the *p*-value.

   The 7% (peanut example) is called a *p*-value. Assuming the company's claim of 50% peanuts in the can is correct, the *p*-value is the probability of getting the results you did (or more extreme) purely by chance.

   A *p*-value less than or equal to 5% is considered statistically significant, to where the researcher would reject the assumption that the observed results were due to random variation and conclude there is strong evidence to support that the results indicate the claim is not true.

   The concept of a *p*-value was formally introduced by Karl Pearson, in his Pearson's Chi-Squared test. The use of the *p*-value in statistics was popularized by Sir Ronald Fisher. In his book *Statistical Methods for Research Workers* (1925), Fisher proposed the level $p = 0.05$ as a possible limit for statistical significance and applied this to a Normal distribution, thus yielding the rule of two standard deviations (on a Normal distribution) for statistical significance using the empirical rule, or 68–95–99.7 rule.

2. Activity to illustrate the general rule that a *p*-value of less than or equal to 5% is considered statistically significant. *Note:* This activity usually takes a few minutes to complete.

   » You will need a deck of all red cards. All the cards need to have the same design on the front so they look like a regular deck of cards. Have the cards in the box so the students assume the box contains the normal arrangement of 26 red and 26 black cards.

   » Tell the students you are going to randomly divide them into two groups based on the color of the card. Red card they are in Group 1 and black card in Group 2.

   » Remove the cards from the box and carefully shuffle them without the students seeing any of the red color on the back of the cards.

   » Go to the first student and turn over a card. Since it will be red, tell the student s/he is in Group 1.

   » Go to the second student and turn over a card. That student will also be in Group 1.

   » Continue until the students get suspicious, usually around the fourth or fifth student.

## Further Exploration and Extensions Cont.

- » Once they get suspicious, stop and discuss the results:
- » Why did you get suspicious?

***Answer:*** *They might say too many reds in a row.*

- » What did they expect to happen? Why did they expect this?

***Answer:*** *Expect about half the cards to be red and the others black. If this were a regular deck of cards, we would expect half to be red and half to be black.*

- » What is the probability that we would get this many red cards in a row, assuming this was a regular deck of cards?

***Answer:*** *4th student – $(1/2)^4$, about 0.06; 5th student – $(1/2)^5$, about 0.03*

Point out that they got suspicious when the probability of observing five red cards in a row was about 3%. They assumed the deck was a regular deck of cards—that is, the probability of turning over a red card was 50% and they reacted (the results they saw were unusual) between 6% and 3%.

*Note:* The process followed is comparable to what is done in traditional hypothesis testing. Hypothesis testing refers to the formal procedures used by statisticians to reject or fail to reject the null hypothesis. In the flipping card example, the null hypothesis is the probability of turning over a red card is 50%. We assumed the null hypothesis (probability of red 50%) is true. Given the results of the card-turning simulation, we decided to reject the null hypothesis and conclude the probability of turning over a red card is not 50%.