# Selecting a Sample

Write down three things you learned about taking a sample from the work we have done today.

1. Would any of the following provide a random sample of letters used in text of the book *Harry Potter and the Sorcerer's Stone* by J.K. Rowling? Explain your reasoning.

   a. Use the first letter of every word of a randomly chosen paragraph.

   b. Number all of the letters in the words in a paragraph of the book, cut out the numbers, and put them in a bag. Then, choose a random set of numbers from the bag to identify which letters you will use.

   c. Have a family member or friend write down a list of their favorite words, and count the number of times each of the letters occurs.

2. Indicate whether the following are random samples from the given population, and explain why or why not.

   a. Population: All students in school; sample includes every fifth student in the hall outside of class.

   b. Population: Students in your class; sample consists of students that have the letter "s" in their last name.

   c. Population: Students in your class; sample selected by putting their names in a hat and drawing the sample from the hat.

   d. Population: People in your neighborhood; sample includes those outside in the neighborhood at 6:00 p.m.

   e. Population: Everyone in a room; sample selected by having everyone toss a coin, and those that result in heads are the sample.

3. Consider the two sample distributions of the number of letters in randomly selected words shown below:

Number of letters in words from two random samples
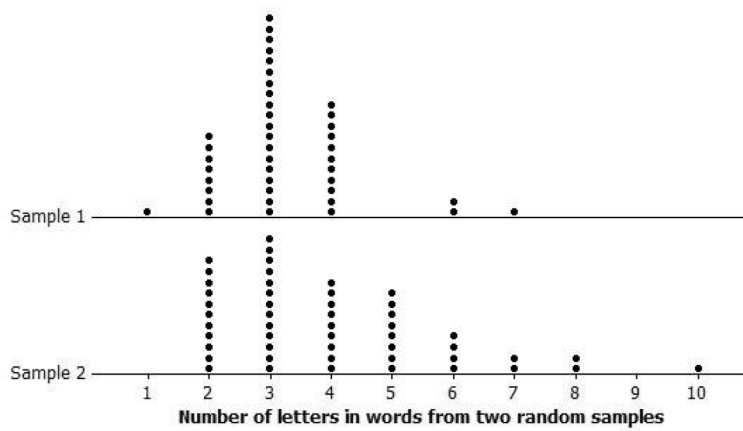
a. Describe each distribution using statistical terms as much as possible.

b. Do you think the two samples came from the same poem? Why or why not?

4. What questions about samples and populations might you want to ask if you saw the following headlines in a newspaper?

a. "Peach Pop is the top flavor according to 8 out of 10 people."

b. "Candidate X looks like a winner! 10 out of 12 people indicate they will vote for Candidate X."

c. "Students overworked. Over half of 400 people surveyed think students spend too many hours on homework."

d. "Action/adventure was selected as the favorite movie type by an overwhelming 75% of those surveyed."

> **Write down three things you learned about taking a sample from the work we have done today.**
>
> *A random sample is one where every element in the set has an equal chance of being selected.*
>
> *When people just choose a sample they think will be random, it will usually be different from a real random sample.*
>
> *Random samples are usually similar to the population.*

The Problem Set is intended to reinforce material from the prior lesson and have students think about examples of samples that are random and those that are not.

> 1. Would any of the following provide a random sample of letters used in text of the book *Harry Potter and the Sorcerer's Stone* by J.K. Rowling? Explain your reasoning.
>
>   a. Use the first letter of every word of a randomly chosen paragraph.
>
>   *This is not a random sample. Some common letters, like "u", don't appear very often as the first letter of a word and may tend to be underrepresented in the sample.*
>
>   b. Number all of the letters in the words in a paragraph of the book, cut out the numbers, and put them in a bag. Then, choose a random set of numbers from the bag to identify which letters you will use.
>
>   *This would give you a random sample of the letters.*
>
>   c. Have a family member or friend write down a list of their favorite words, and count the number of times each of the letters occurs.
>
>   *This would not be a random sample. They might like words that rhyme or that all start with the same letter. The list might also include words not in the book.*
>
> 2. Indicate whether the following are random samples from the given population, and explain why or why not.
>
>   a. Population: All students in school; sample includes every fifth student in the hall outside of class.
>
>   *Sample response: No, because not everyone in school would be in our hall before class—our hall only has sixth graders in it, so the seventh and eighth graders would not have a chance to be chosen.*
>
>   b. Population: Students in your class; sample consists of students that have the letter "s" in their last name.
>
>   *Sample response: No, because students that do not have the letter "s" in their last name would not have a chance to be chosen.*
>
>   c. Population: Students in your class; sample selected by putting their names in a hat and drawing the sample from the hat.
>
>   *Sample response: Yes, everyone would have the same chance to be chosen.*
>
>   d. Population: People in your neighborhood; sample includes those outside in the neighborhood at 6:00 p.m.
>
>   *Sample response: No, because people who are not in the neighborhood at that time have no chance of being selected.*

e. Population: Everyone in a room; sample selected by having everyone toss a coin, and those that result in heads are the sample.

Sample response: Yes, everyone would have the same chance to be chosen.

3. Consider the two sample distributions of the number of letters in randomly selected words shown below:
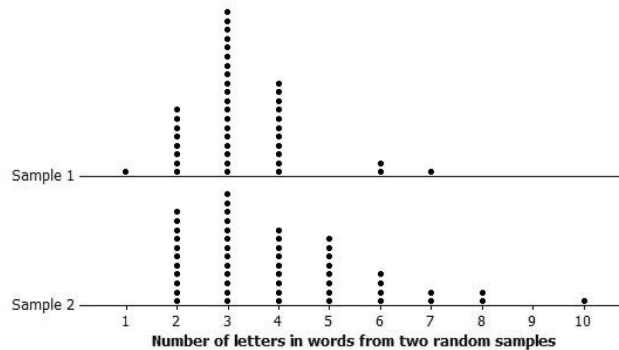
a. Describe each distribution using statistical terms as much as possible.

Answers will vary; the top distribution seems to have both a median and balance point, or mean, at 3, with a minimum of 1 letter in a word and a maximum of 7 letters. Most of the words in the sample were 2 to 4 letters long. The bottom distribution seems more skewed with the median of about 4 letters. The smallest number of letters was 2, and the largest was 10 letters. Most of the letters in this sample had between 2 and 5 letters.

b. Do you think the two samples came from the same poem? Why or why not?

Sample response: The samples could have come from the same poem, but the distributions seem different both with respect to shape and to a measure of center, so it seems more likely that they were from two different populations.

4. What questions about samples and populations might you want to ask if you saw the following headlines in a newspaper?

a. "Peach Pop is the top flavor according to 8 out of 10 people."

Sample response: How were the people selected? How many people were surveyed? What were the choices, and how many did not like Peach Pop?

b. "Candidate X looks like a winner! 10 out of 12 people indicate they will vote for Candidate X."

Sample response: How was the sample chosen? Were the people selected at random, or were they friends of Candidate X?

c. "Students overworked. Over half of 400 people surveyed think students spend too many hours on homework."

Sample response: Who was surveyed, and how were they selected? Was the survey given to students in a school?

d. "Action/adventure was selected as the favorite movie type by an overwhelming 75% of those surveyed."

Sample response: Was the survey given at a movie theater showing an action/adventure movie where people were there because they like that kind of movie?