Lesson 5: Discrete Random Variables

Classwork

Example 1: Types of Data

Recall that the sample space of a chance experiment is the set of all possible outcomes for the experiment. For example, the sample space of the chance experiment that consists of randomly selecting one of ten apartments in a small building would be a set consisting of the ten different apartments that might have been selected. Suppose that the apartments are numbered from 1 to 10. The sample space for this experiment is {1,2,3,4,5,6,7,8,9,10}.

Cards with information about these ten apartments will be provided by your teacher. Mix these cards and then select one. Record the following information for the apartment you selected.

Number of bedrooms: Floor number:

Size (sq. ft.):   Distance to elevator:

Color of walls: Floor type:

Exercise 1

1. Sort the features of each apartment into three categories:

   a. Describe how the features listed in each category are similar.

b. A random variable associates a number with each outcome of a chance experiment. Which of the features are random variables? Explain.

Example 2: Random Variables

One way you might have sorted these variables is whether they are based on counting (such as the number of languages spoken) or based on measuring (such as the length of a leaf). Random variables are classified into two main types: *discrete* and *continuous*.

- A *discrete random variable* is one that has possible values that are isolated points along the number line. Often, discrete random variables involve counting.

- A *continuous random variable* is one that has possible values that form an entire interval along the number line. Often, continuous random variables involve measuring.

Exercises 2–3

2. Choose six variables from the list in Exercise 1. For each one, give a specific example of a possible value the variable might have taken on, and identify the variable as discrete or continuous.

3. Suppose you were collecting data about dogs. Give at least two examples of discrete and two examples of continuous data you might collect.

Exercises 4–8: Music Genres

People like different genres of music: country, rock, hip hop, jazz, etc. Suppose you were to give a survey to people asking them how many different music genres they like.

4.   What do you think the possible responses might be?

5.   The table below shows 11,565 responses to the survey question: How many music genres do you like listening to?

Table 1: Number of music genres survey responders like listening to

Number of music genres	0	1	2	3	4	5	6	7	8
Number of responses	568	2,012	1,483	654	749	1,321	1,233	608	2,937

Find the relative frequency for each possible response (each possible value for number of music genres), rounded to the nearest hundredth. (The relative frequency is the proportion of the observations that take on a particular value. For example, the relative frequency for 0 is $\frac{568}{11565}$.)

6.   Consider the chance experiment of selecting a person at random from the people who responded to this survey. The table you generated in Exercise 5 displays the probability distribution for the random variable number of music genres liked. Your table shows the different possible values of this variable and the probability of observing each value.

a.   Is the random variable discrete or continuous?

b.   What is the probability that a randomly selected person who responded to the survey said that they like 3 different music genres?

c.　Which of the possible values of this variable has the greatest probability of being observed?

d.　What is the probability that a randomly selected person who responded to the survey said that they liked 1 or fewer different genres?

e.　What is the sum of the probabilities of all of the possible outcomes? Explain why your answer is reasonable for the situation.

7.　The survey data for people age 60 or older are displayed in the graphs below.
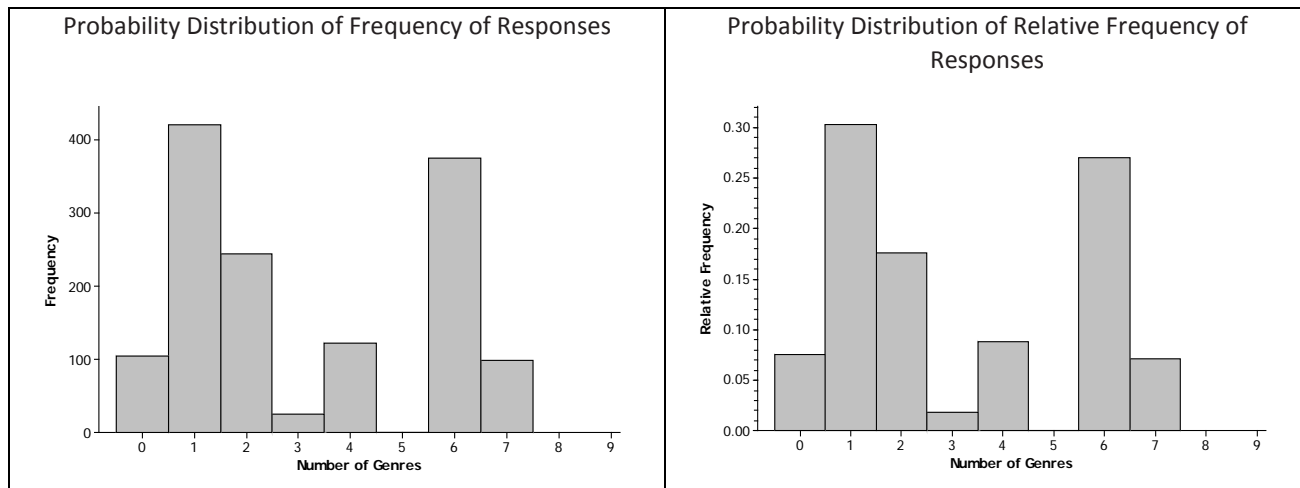
a.　What is the difference between the two graphs?

b.　What is the probability that a randomly selected person from this group of people age 60 or older chose 4 music genres?

c. Which of the possible values of this variable has the greatest probability of occurring?

d. What is the probability that a randomly selected person from this group of people age 60 or older chose 5 different genres?

e. Make a conjecture about the sum of the relative frequencies. Then check your conjecture using the values in the table.

8. Below are graphs of the probability distribution based on responses to the original survey and based on responses from those age 60 and older.

Probability Distribution from Original Survey

Probability Distribution for 60 Years and Older
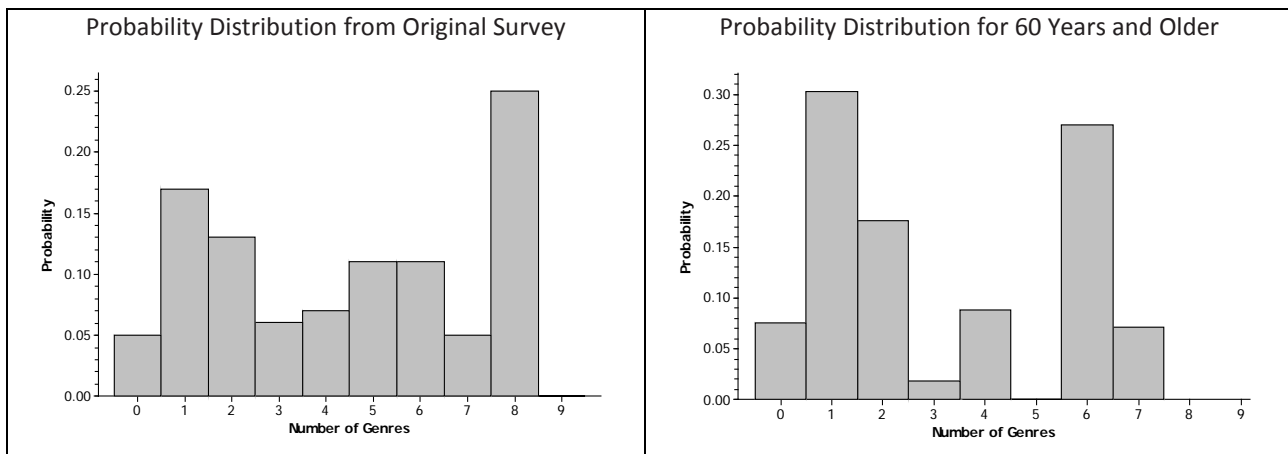
Identify which of the statements are true and which are false. Give a reason for each claim.

a. The probability that a randomly selected person chooses 0 genres is greater for those age 60 and older than for the group that responded to the original survey.

b. The probability that a randomly selected person chooses fewer than 3 genres is smaller for those age 60 and older than for the group that responded to the original survey.

c. The sum of the probabilities for all of the possible outcomes is larger for those age 60 and over than for the group that responded to the original survey.

Exercises 9–11: Family Sizes

The table below displays the distribution of the number of people living in a household according to a recent U.S. Census. This table can be thought of as the probability distribution for the random variable that consists of recording the number of people living in a randomly selected U.S. household. Notice that the table specifies the possible values of the variable, and the relative frequencies can be interpreted as the probability of each of the possible values.

Table 2: Relative frequency of the number of people living in a household

Number of People	Relative Frequency
1	0.24
2	0.32
3	0.17
4	0.16
5	0.07
6	0.02
7 or more	0.02

9. What is the random variable, and is it continuous or discrete? What values can it take on?

10. Use the table to answer each of the following.

   a. What is the probability that a randomly selected household would have 5 or more people living there?

b. What is the probability that 1or more people live in a household? How does the table support your answer?

c. What is the probability that a randomly selected household would have fewer than 6 people living there? Find your answer in two different ways.

11. The probability distributions for the number of people per household in 1790, 1890, and 1990 are below.

Number of people per household	1	2	3	4	5	6	7 or more
1790: Probability	0.03	0.08	0.12	0.14	0.14	0.13	0.36
1890: Probability	0.04	0.13	0.17	0.17	0.15	0.12	0.23
1990: Probability	0.24	0.32	0.17	0.16	0.07	0.02	0.01

Source: U.S. Census Bureau (www.census.gov)

a. Describe the change in the probability distribution of the number of people living in a randomly selected household over the years.

b. What are some factors that might explain the shift?

Lesson Summary

- Random variables can be classified into two types: discrete and continuous.

- A discrete random variable is one that has possible values that are isolated points along the number line. Often, discrete random variables involve counting.

- A continuous random variable is one that has possible values that form an entire interval along the number line. Often, continuous random variables involve measuring.

- Each of the possible values can be assigned a probability and the sum of those probabilities is 1.

- Discrete probability distributions can be displayed graphically or in a table.

Problem Set

1. Each person in a large group of children with cell phones was asked, "How old were you when you first received a cell phone?"

   The responses are summarized in the table below.

Age in years	Probability
9	0.03
10	0.06
11	0.11
12	0.23
13	0.23
14	0.14
15	0.11
16	0.08
17	0.01

   a. Make a graph of the probability distribution.

   b. The bar centered at 12 in your graph represents the probability that a randomly selected person in this group responded that they first received a cell phone at age 12. What is the area of the bar representing age 12? How does this compare to the probability corresponding to 12 in the table?

   c. What do you think the sum of the areas of all of the bars will be? Explain your reasoning.

   d. What is the probability that a randomly selected person from this group first received a cell phone at age 12 or 13?

   e. Is the probability that a randomly selected person from this group first received a cell phone at an age older than 15 greater than or less than the probability that a randomly selected person from this group first received a cell phone at an age younger than 12?

2. The following table represents a discrete probability distribution for a random variable. Fill in the missing values so that the results make sense; then, answer the questions.

Possible value	4	5	10	12	15
Probability	0.08	???	0.32	0.27	???

   a. What is the probability that this random variable takes on a value of 4 or 5?

   b. What is the probability that the value of the random variable is not 15?

   c. Which possible value is least likely?

3. Identify the following as true or false. For those that are false, explain why they are false.

   a. The probability of any possible value in a discrete random probability distribution is always greater than or equal to 0 and less than or equal to 1.

   b. The sum of the probabilities in a discrete random probability distribution varies from distribution to distribution.

   c. The total number of times someone has moved is a discrete random variable.

4. Suppose you plan to collect data on your classmates. Identify three discrete random variables and three continuous random variables you might observe.

5. Which of the following are not possible for the probability distribution of a discrete random variable? For each one you identify, explain why it is not a legitimate probability distribution.

Possible value	1	2	3	4	5
Probability	0.1	0.4	0.3	0.2	0.2

Possible value	1	2	3	4
Probability	0.8	0.2	0.3	−0.2

Possible value	1	2	3	4	5
Probability	0.2	0.2	0.2	0.2	0.2

6. Suppose that a fair coin is tossed 2 times, and the result of each toss (H or T) is recorded.

   a. What is the sample space for this chance experiment?

   b. For this chance experiment, give the probability distribution for the random variable of total number of heads observed.

7. Suppose that a fair coin is tossed 3 times.

    a. How are the possible values of the random variable of total number of heads observed different from the possible values in the probability distribution of Problem 6(b)?

    b. Is the probability of observing a total of 2 heads greater when the coin is tossed 2 times or when the coin is tossed 3 times? Justify your answer.