CHAPTER 10 PROJECT

Benford's Law: What Do Electricity Bills, House Prices, Population Numbers, Death Rates, and the Lengths of Rivers Have in Common?

Benford's law, also known as the Newcomb–Benford law, states that the first digit in real-life data is more likely to be a small number (such as 1) than it is to be a large number (such as 9). Simon Newcomb first observed the phenomenon in 1881 (with Frank Benford independently rediscovering it in 1938) after testing data from 20 different fields of study, including the surface areas of rivers, physical constants, and molecular weights. In this project, you will explore some applications of Benford's law.

Suppose we wrote each 4-digit number on separate pieces of paper and put them in a hat. (Notice that the first digit must be 1, 2, 3, 4, 5, 6, 7, 8, or 9 while the remaining digits could also include 0.)

1. How many numbers in the hat start with the digit 1?

2. What is the probability that a randomly selected number from the hat has 1 as its first digit?

3. Suppose the hat contained every 12-digit number. What would be the probability of randomly selecting a number with the first digit equal to 1?

Now, let's see if switching the first digit makes any difference in the probability.

4. How many 4-digit numbers in the hat start with the digit 9?

5. What is the probability that a randomly selected number from the hat has 9 as its first digit?

6. Suppose the hat contained every 12-digit number. What would be the probability of randomly selecting a number with the first digit equal to 9?

The investigation in parts 1 through 6 illustrate what we call a *uniform* distribution. It seems like a reasonable thing to accept; there shouldn't be any reason why 1 would be more or less likely to occur as the first digit of a randomly generated number than 9.

7. Suppose you are given a list of 1000 numbers. If we assume there is a uniform distribution of numbers within the list, how many of those numbers would begin with 1?

Benford's law, strangely enough, states that the distribution of first digits is indeed not uniform but it follows the probabilities in the following tables.

Digit	Probability of Being the First Digit
1	30.1%
2	17.6%
3	12.5%
4	9.7%
5	7.9%
6	6.7%
7	5.8%
8	5.1%
9	4.6%

8. Suppose that you were presented with a set of 1000 numbers. Use Benford's law to calculate how many numbers would you expect to start with 1.

9. Use Benford's law to calculate how many numbers would you expect to start with 9.

10. Compare the values found in parts 8 and 9. Explain what the values mean.

Let's test Benford's law using a few examples.

11. Write a list of the first 30 powers of 2. That is, write the numbers that represent 21, 22, 23, …, 230. (Hint: The use of Microsoft Excel or a calculator will help with the calculations!)

12. How many numbers in the first 30 powers of 2 begin with the number 1?

13. What is the percentage of numbers that start with the digit 1 in the first 30 powers of 2?

14. What is the probability that a randomly selected power of 2 starts with the digit 1 if it is chosen from the list of the first 30 powers of 2? Does this result align with Benford's Law?

15. Next, perform an internet search to find a list of counties in South Carolina with a population under 250,000. What is the percentage of counties whose population starts with the digit 1? How does this compare to Benford's law?

16. Perform an internet search to learn about different ways Benford's law can be used to detect fraud. Describe one instance where Bendford's law detected fraud.