Lecture Examples using Simulation-Based Inference and Active Learning

Jimmy A. Doi

ABSTRACT

This article provides an overview of a statistics lecture given in summer 2017 at Takasaki Super Science High School located in Gunma, Japan. The lecture consisted of four topics: simulation-based inference (SBI), Benford's Law, longest run of heads or tails, and the Chaos Game. Each of the topics required active participation by the students. Lectures using SBI and active learning methods can be very effective to help students better understand statistical and mathematical concepts. A complete description of each of the four lecture topics is provided including many relevant resources.

Keywords: Statistics education, simulation-based inference, active learning, curriculum, applets, Shiny

1. Introduction

In summer 2017 I was invited to give a statistics lecture at Takasaki Super Science High School in Gunma, Japan. Although I was given the option to cover topics of my choosing, I was asked to reserve one topic for simulation-based inference, which is a method that has been growing in popularity for teaching statistical inference.

The lecture was 2-hours in length and, to make the lecture more understandable to the students, I was asked by the high school staff to give the presentation entirely in Japanese. The lecture topics I covered were simulation-based inference, Benford's Law, longest run of heads or tails, and the Chaos Game. Each of the four topics contained some active learning component where students actively participated during the lecture discussion. For each topic I used freely available web-based applications for demonstration purposes. Complete details for these applications will be provided later.

Section 2 is devoted to the simulation-based inference discussion. The next lecture topic, Benford's Law, is discussed in Section 3. This is followed in Section 4 by the longest run of heads or tails lecture topic. The last lecture topic on the Chaos Game is discussed in Section 5. Section 6 is devoted to the lecture summary and Section 7 includes concluding remarks.

During the lecture, I led the students through a set of course notes that they wrote in during the presentation. For the interested reader, the student version of the course notes (without answers) can be found at https://bit.ly/Takasaki-notes-student. The instructor version of the course notes (with answers) can be found at https://bit.ly/Takasaki-notes-instructor. Although most of the course notes were written in English, during the lecture I translated and explained each page in Japanese.

2. Topic #1: Simulation-Based Inference (SBI)

The first topic I discussed uses an example from simulation-based inference (SBI). Before I provide an overview of the lecture discussion, I would like to give some background information about SBI and provide SBI references for the interested reader.

(1) Introduction to SBI

As the name suggests, SBI is an approach that exposes students to the logic of statistical inference via simulations. In a traditional setting statistical inference is taught by first introducing formal probability and sampling distributions. With SBI we can use simple devices such as dice, coins, cards, and computer applications to perform simulations and generate approximate sampling distributions. With this approach students can quickly gain an understanding about statistical inference even if they have a minimal statistical or mathematical background. In fact, some instructors even use SBI examples on the first day of an introductory statistics class (see Roy et al. 2014). Although none of the Takasaki High School students had any prior statistics knowledge they all seemed to understand the SBI lecture example very well.

Tintle et al. (2011) claim that SBI addresses at least two major criticisms of the traditional curriculum. First, SBI allows students to focus on the logic of inference instead of the derivation of theoretical distributions, which may not seem well-connected to inference. Second, SBI allows students to be exposed to modern computational methods that are growing in popularity.

Recently there has been an increasing use of SBI as an approach for teaching introductory statistics courses (Tintle et al. 2015; Tintle et al. 2011). In addition, when compared to students exposed to the traditional curriculum, studies have shown that there is improved conceptual understanding among students in courses that use SBI (Chance and McGaughey 2014; Hildreth et al. 2018; Tintle et al. 2014).

As mentioned previously, a key component of SBI is the use of computer simulations. Computer simulations can greatly enhance the understanding of concepts such as randomness, sampling, and variability (Chance et al. 2007; delMas et al. 1999; Garfield and Ben-Zvi, 2008). The use of technology to introduce and reinforce essential concepts has also been promoted in the Guidelines for Assessment and Instruction in Statistics Education (GAISE) for introductory statistics at the college level (Franklin and Garfield, 2006; ASA, 2016) as well as for teaching statistics at the secondary and primary levels (ASA, 2007). For each of the four topics I discussed in class, I used web-based applications for demonstration purposes. For the SBI lecture I used a Java applet from the *Rossman/Chance Applet Collection* (http://www.rossmanchance.com/applets). For the remaining three topics, I used web-based applications known as Shiny apps which I will introduce later.

Some textbooks that use SBI have been written by Chance and Rossman (2018), Diez et al. (2014), Lock et al. (2017), and Tintle et al. (2016). A recent high school level textbook that uses SBI has been written by Tintle et al. (2018). Finally, an SBI blog containing many resources can be found at https://www.causeweb.org/sbi.

(2) Motivating Example

For the first lecture topic I began by describing an experiment conducted by researchers at Yale University. These researchers wondered whether children less than a year old can recognize the difference between helper, friendly behavior as opposed to mean, unhelpful behavior. In order for the students to better understand the experimental setting, I showed a YouTube video created by the researchers (https://youtu.be/anCaGBsBOxM) containing details of one of the experiments they conducted. Complete information about the study can be found in Hamlin et al. (2007).

In their experiment sixteen 10-month old infants participated. Each infant was shown a stage that showed a “climber” character who tried to repeatedly climb a hill but was unable to succeed. The climber character is shown in Figure 1(a). Then each infant was shown two scenes. In one scene, a “helper” character appeared from the bottom of the hill and helped the climber character reach the top. The helper character is shown in Figure 1(b). In the other scene, a “hinderer” character appeared from the top of the hill and pushed the climber character to the bottom. The hinderer character is shown in Figure

1(c). After alternating these two scenes multiple times, the infant was presented with the helper and hinderer characters and the researchers recorded which character the infant reached for first. Figure 1(d) shows an infant reaching for one of the characters. The experimental outcome was that 14 out of 16 infants selected the helper character.

It is important to note that the researchers randomized the color and shape of the climber, helper, and hinderer characters when repeating the experiment. Also, when the infant was shown the two characters at the end of the experiment, the researchers randomized which character was placed to the left and to the right of the infant. In this way the researchers were able to eliminate color, shape, and location preference as reasons for selecting a character.

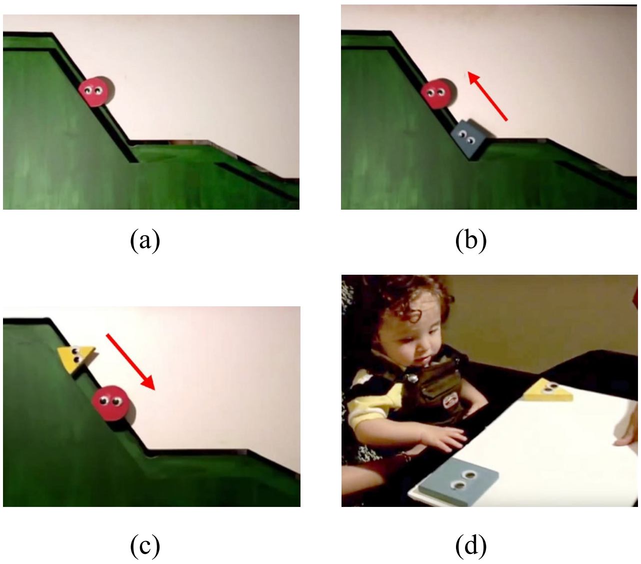


Figure 1. (a-d). Screenshots from the YouTube video describing the infant experiment conducted by Hamlin et al. (2007).

(3) Class Activity

I then led the students through a class activity concerning the experiment and its outcome. A photo taken during this activity is shown in Figure 2. Given that 14 out of 16 infants chose the helper character, I asked for possible explanations for this result. One obvious explanation is that the infants prefer the helper character. This would be the researchers' conjecture. Another explanation is that the infants had no preference for either character and that the result occurred by chance. Assuming infants had no preference for either character, we needed to investigate whether 14 out of 16 choosing the helper character would be unusual. I then asked the students to describe how we could use a common device to simulate the infants' selection process assuming they had no preference for either character. The students stated that a coin toss would be perfectly suited for this purpose.

Figure 2. Interaction with Takasaki High School students during the SBI lecture discussion.

After distributing coins to everyone in class, I asked each student to toss their coin 16 times. I explained that each toss would simulate the selection process of an infant having no character preference. If the outcome was a "head" then this would mean the helper character was selected. If the outcome was a "tail" then this would mean the hinderer character was selected. I asked each student to record the number of times head occurred in their 16 tosses. For their simulation, this value would correspond to the number of infants who selected the helper character.

I then asked all students to report the number of times the helper character was selected based on their simulations. I summarized their results on the chalkboard by displaying the corresponding dotplot which I have reproduced in Figure 3. Given this dotplot, I asked the students whether the result obtained by the researchers (helper selected 14 times) would be surprising assuming the infants had no preference. I also asked whether, based on this plot, the students would feel there was sufficient evidence to conclude that the infants did in fact prefer the helper character. Looking at the plot, the students agreed that the researchers' result appeared to be unusual, if indeed, babies had no preference, and the students felt confident that infants in general preferred the helper character.

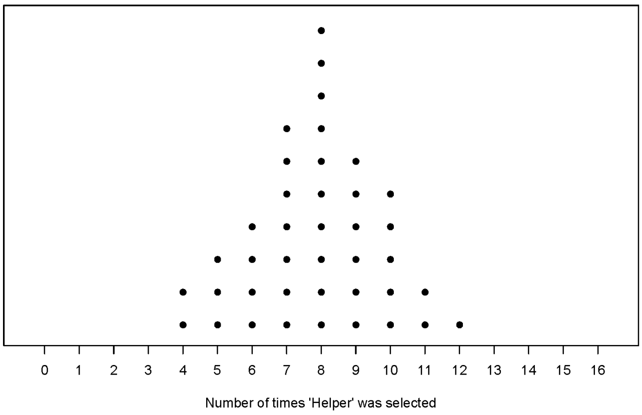


Figure 3. Dotplot of the number of times the 'helper' character was selected from the students' coin toss simulations.

(4) Java Applet for Simulation

The dotplot we created provided an indication of how unusual the researchers' result was assuming the infants had no character preference, but I explained to the students that the plot is an incomplete picture as it is based on only 40 simulation results (one for each student). Ideally, we would like to simulate the infants' selection process thousands and thousands of times. To investigate this, I introduced a Java applet that simulates coin tosses. This applet can be found at http://www.rossmanchance.com/applets/OneProp/OnePropJPN.html.

A screenshot of this applet is shown in Figure 4. Using the applet I demonstrated for the students how to simulate 16 tosses of a fair coin. By modifying the number of repetitions, I showed how to quickly generate 100,000 simulations. Figure 4 shows the results of 100,000 simulations and the corresponding dotplot of the number of heads. Finally, I described how to use the applet to determine the proportion of all trials that gave results at least as extreme as what was observed in the researchers' original experiment. For the simulation results shown in Figure 4, this proportion was 0.0021.

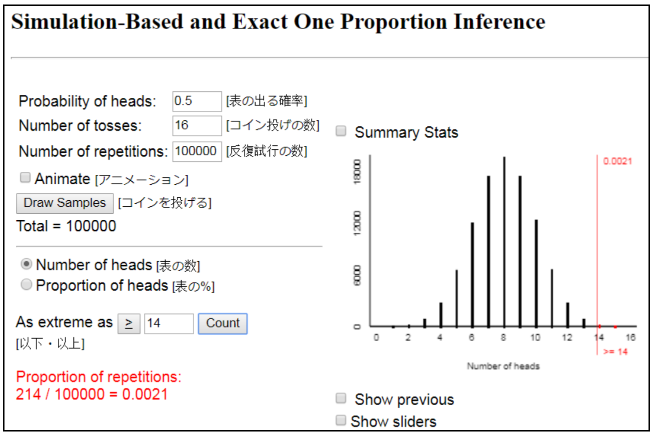


Figure 4. Coin toss simulation Java applet screenshot. Display shows the distribution of the number of heads from 100,000 simulations of 16 tosses. The display also reports the proportion of outcomes that were at least as extreme as 14 heads.

Given that the lecture was presented in a computing lab, I asked all students to practice using the applet on their laptops by simulating several repetitions of the experiment. A photo of a student using the applet on a laptop computer is shown in Figure 5. After each student performed 40 repetitions, I asked them to compare their applet's dot plot to the dotplot we had created earlier in class (shown in Figure 3). Each student noticed that the images were similar.

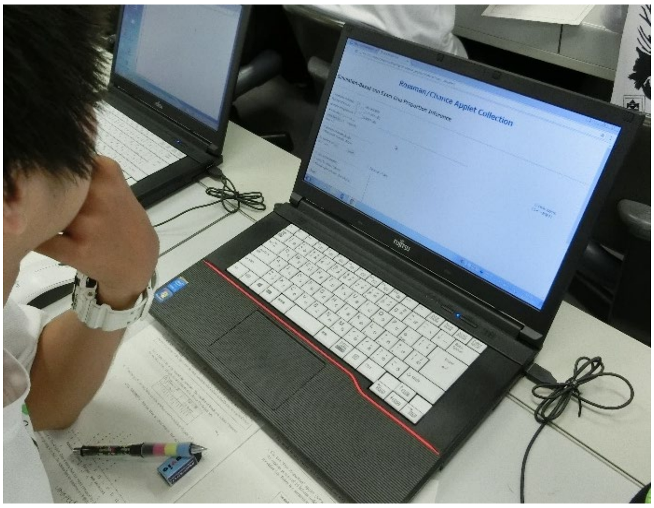


Figure 5. A student using the Java applet simulation.

After students had gained sufficient practice, I asked them to reset the applet and perform 100,000 simulations. I asked the students to notice that their applet's dotplot of the number of heads looks almost identical to the dotplot from my earlier demonstration (shown in Figure 4). I explained the reason for this is because, as the number of repetitions increases, the simulation distribution converges to the binomial distribution based on 16 trials and success probability equal to 0.5, written as $B(16, 0.5)$. The students had studied the binomial distribution during a previous lecture on probability and so they were familiar with this distribution.

I then asked the students to use the applet to determine the proportion of simulation outcomes that were at least as extreme as the researchers' outcome. It was not the case that everyone had the exact same proportions, however their values were all very close to 0.002. I explained the reason for this agreement was again due to the fact that their simulation distributions were approximately $B(16, 0.5)$. Using the $B(16, 0.5)$ distribution I asked the students to compute by hand the exact proportion of an outcome at least as extreme as 14 and they were able to confirm the answer is 0.00209. At this point they understood why everyone's proportions were very close to 0.002.

Seeing how small this proportion was the students agreed once again that the researchers' result was very unusual assuming the infants had no character preference and that their simulation results provided sufficient evidence the infants preferred the helper character.

(5) SBI and Traditional Curriculum

It is important to point out that SBI methods need not replace traditional teaching content and methods but can instead *enhance* them. In a traditional setting we usually teach statistical inference by introducing the following concepts: null and alternative hypotheses, test statistic, null distribution, and P-value. In many of the introductory statistics classes I teach at my university, I of course introduce and cover all these topics. But recently, for some of these introductory classes, I have initially introduced this SBI example involving the infant experiment. Then, at a later point when I formally introduce these inferential concepts mentioned previously, I remind my students about the SBI example and, by doing so, I believe the students are better able to comprehend the inferential concepts.

For example, when I formally introduce "null hypothesis," after reminding them about the experiment, students immediately determine "infants have no character preference" as the natural null hypothesis in that setting. For "alternative hypothesis," students suggest "infants prefer the helper character" as the natural alternative hypothesis. For "test statistic," students will recall that we counted the number of heads from the coin tossing simulation and this number could be naturally thought of as the test statistic. For "null distribution," keeping in mind the no character preference as the null hypothesis, the test statistic follows the binomial distribution, which is also the null distribution. Students will recall that, as the number of repetitions increases, the simulation distribution from the applet converges to this binomial distribution. With this in mind, students can have an intuitive understanding of what the null distribution is. Finally, for "P-value," students will recall that we determined the proportion of simulation outcomes that were at least as extreme as the observed experimental result. As the number of repetitions increases, the simulation proportion converges to the P-value, and so students can have an intuitive understanding of what the P-value is. And so, by first introducing an SBI example, I believe it is possible for students to better comprehend inferential reasoning. Therefore, I strongly believe that SBI methods can greatly enhance traditional teaching content and methods.

3. Topic #2: Benford's Law

(1) Motivating Example

For the second lecture topic I began by having the students examine the leading digits of the first 18 values of 2^n (2^1 to 2^{18}). These values are shown in Table 1 with leading digits underlined.

For a given set of numbers, the possible values of the leading digit are 1, 2, …, 9. I asked the students whether,

Table 1 Leading digits for 2^1 to 2^{18}.

2^1 = 2	2^7 = 128	2^{13} = 8192
2^2 = 4	2^8 = 256	2^{14} = 16384
2^3 = 8	2^9 = 512	2^{15} = 32768
2^4 = 16	2^{10} = 1024	2^{16} = 65536
2^5 = 32	2^{11} = 2048	2^{17} = 131072
2^6 = 64	2^{12} = 4096	2^{18} = 262144

in general, they would expect any leading digits to occur more frequently than others. The students replied that they expected the values to occur uniformly, that is with equal probability 1/9. Applying this assumption to the first 18 values of 2^n, this would imply each possible leading digit would occur roughly twice. Returning to the list of 18 values I asked the students to determine the observed frequencies of leading digits and the students quickly found that these frequencies were not close to uniform. The comparison of uniform frequencies and observed frequencies of leading digits for 2^1 to 2^{18} are shown in Table 2. Most students immediately noted the curious pattern where the observed frequencies seem to *decrease* as the leading digit value increases.

Table 2 Uniform versus observed frequencies of leading digits for 2^1 to 2^{18}.

	Leading Digit								
	1	2	3	4	5	6	7	8	9
Unif. Freq.	2	2	2	2	2	2	2	2	2
Obs. Freq.	5	4	2	2	1	2	0	2	0

One might argue that observed frequencies would be more uniform if we examined a larger number of values of 2^n. To investigate this, I displayed for the students the first 90 values of 2^n. Having the students work in teams I again asked the students to determine the observed frequencies of the leading digits for these 90 values. With a uniform probability assumption each leading digit would occur 10 times. However, the students determined that the observed frequencies are distinctly non-uniform. The comparison of uniform frequencies and observed frequencies of leading digits for 2^1 to 2^{90} are shown in Table 3. Again, the students noted the unexpected pattern where the observed frequencies tend to *decrease* as the leading digit value increases.

Table 3 Uniform versus observed frequencies of leading digits for 2^1 to 2^{90}.

	Leading Digit								
	1	2	3	4	5	6	7	8	9
Unif. Freq.	10	10	10	10	10	10	10	10	10
Obs. Freq.	27	16	11	9	7	6	5	5	4

A helpful visual aid for the frequencies of leading digits for 2^n is shown in Figure 6. Here the leading digits for 2^1 to 2^{234} are shown in an array. For example, in the first row of the array, the initial values 2, 4, 8, 1, and 3 correspond to the leading digits of 2^1, 2^2, 2^3, 2^4, and 2^5 respectively. I displayed for the students an updated array where all leading digits of 1 are emphasized and circled. This allowed the students to easily see the high frequency of the leading digit 1. This is shown in Figure 6(a). I then showed the students updated arrays marking the locations of the leading digits of 2, 3, 4, …, and 9. With each subsequent updated array the decreasing frequency of the leading digit becomes immediately obvious to the students. The arrays for the leading digits of 3, 6, and 9 are shown in Figures 6(b), 6(c), and 6(d) respectively.

(2) Benford's Law

At this point of the lecture the students were quite intrigued by the unusual pattern of observed frequencies and I then introduced Benford's Law. This law states that the relative frequency of the leading digit (d) of a set of numbers is often determined by the following probability function:

$$P(d) = \log_{10}(1 + 1/d) \qquad \text{where } d = 1, 2, \ldots, 9.$$

The leading digits and corresponding probabilities from Benford's Law are shown in Table 4. With this table the students were able to note the decreasing nature of the probabilities as d increases.

I then displayed a graph comparing the probabilities based on the uniform distribution and Benford's Law.

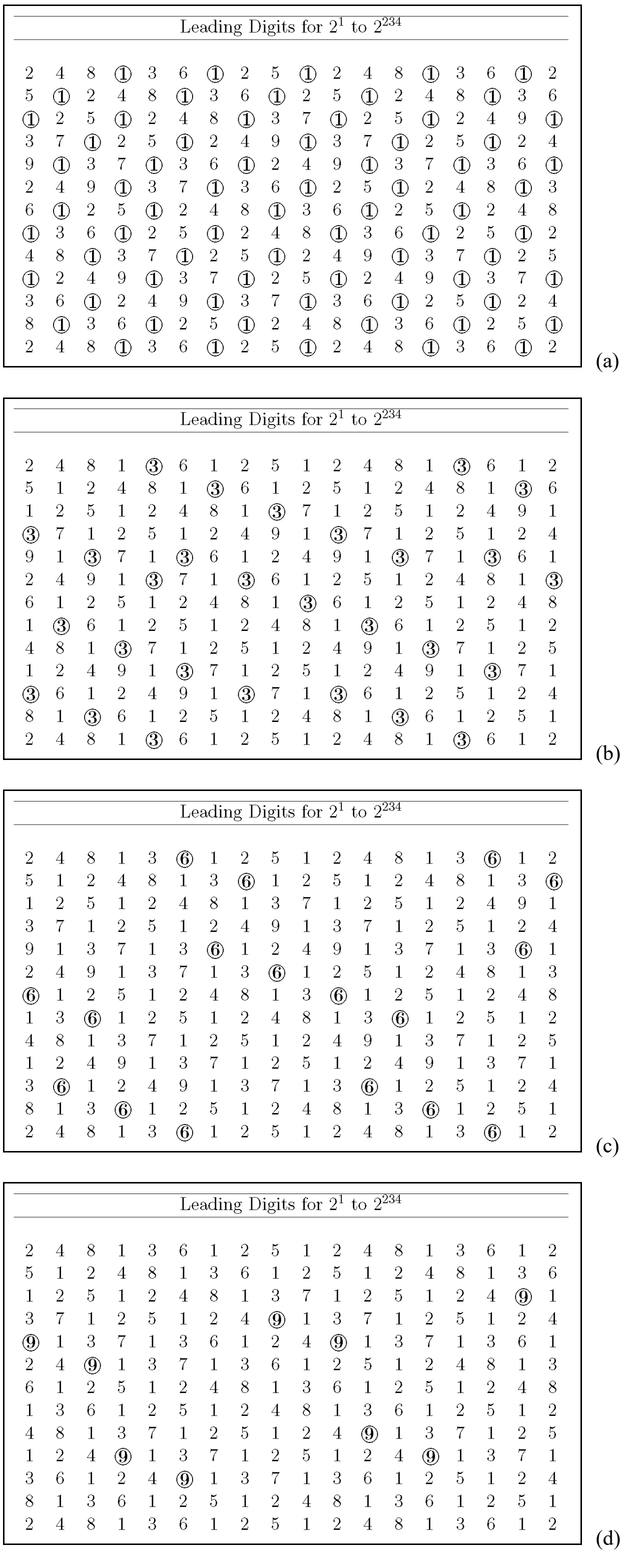


Figure 6. (a-d). Arrays of leading digits for 2^1 to 2^{234} with marked locations for (a) 1, (b) 3, (c) 6, and (d) 9.

With this image students can quickly see the stark contrast between the two probability models. This image is shown in Figure 7(a).

Table 4　Leading digit probabilities based on Benford's Law.

d	$P(d)$
1	$\log_{10}(1+1/1)$ ≒ .301
2	$\log_{10}(1+1/2)$ ≒ .176
3	$\log_{10}(1+1/3)$ ≒ .125
4	$\log_{10}(1+1/4)$ ≒ .097
5	$\log_{10}(1+1/5)$ ≒ .079
6	$\log_{10}(1+1/6)$ ≒ .067
7	$\log_{10}(1+1/7)$ ≒ .058
8	$\log_{10}(1+1/8)$ ≒ .051
9	$\log_{10}(1+1/9)$ ≒ .046

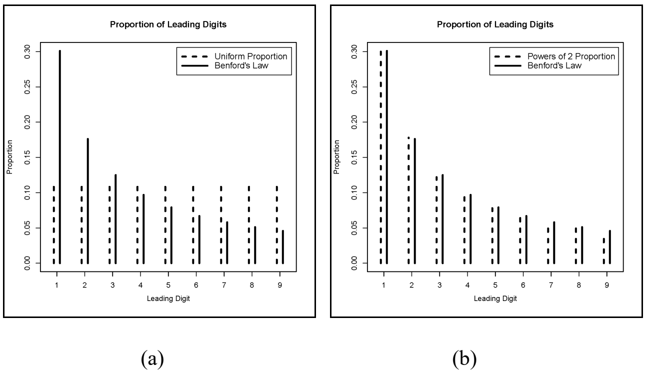


Figure 7. (a) Comparison of leading digit probabilities based on uniform distribution and Benford's Law. (b) Comparison of observed leading digit percentages for 2^1 to 2^{90} and Benford's Law percentages.

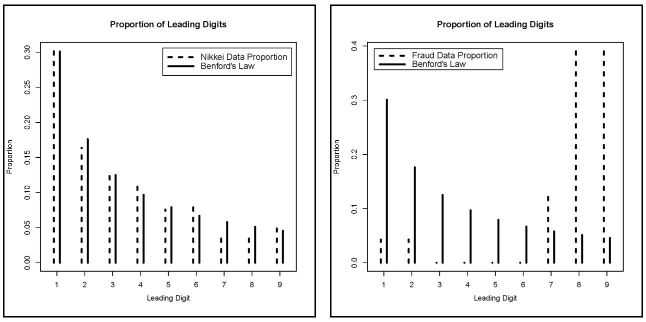
For the earlier class activity where students determined the observed frequencies of leading digits for 2^1 to 2^{90}, I converted the frequencies to percentages and displayed an image comparing these observed percentages to the percentages from Benford's Law. This image is shown in Figure 7(b). With this image the students can see the surprising agreement of observed percentages with Benford's Law across all leading digits.

As n increases, the observed percentages of leading digits for 2^1 to 2^n converge with the percentages from Benford's Law. The agreement of the leading digit distribution with Benford's Law is not unique to the sequence of the powers of 2. The leading digits for the sequence of powers of other bases (3^n, 4^n, 5^n, …) will yield similar behavior. Benford's Law also applies to the leading digits of additive sequences, such as the Fibonacci sequence. Finally, as described by Luque and

Lacasa (2009), a variation of Benford's Law applies to the leading digits of the sequence of prime numbers.

(3) Real-World Examples

There are many real-world data sets where the leading digit distribution is well-described by Benford's Law. Some examples I discussed included the surface areas of lakes in California, Japanese population data from the 2015 Japanese Census, and stock prices from the Nikkei 225 Stock Exchange. For all these examples the leading digit distribution closely follows Benford's Law. A comparison of observed leading digit percentages for the Japanese stock exchange data and Benford's Law percentages is shown in Figure 8(a).



(a) (b)

Figure 8. (a) Comparison of observed leading digit percentages for closing prices of Nikkei 225 stocks and Benford's Law percentages. (b) Comparison of observed leading digit percentages for fraudulent charges and Benford's Law percentages.

The last real-world data example I discussed was from the area of fraud detection. This example, as described by Nigrini (1999), is based on 23 fraudulent checks by an employee in the office of the Arizona State Treasurer. As with stock prices, general payments can have a leading digit distribution that closely follows Benford's Law. I asked the students to examine the 23 fraudulent payments and they immediately found that its leading digit distribution was very different from Benford's Law. There was a high frequency of the leading digit being 8 or 9, and there was a low frequency of the leading digit being 1 or 2. Such behavior is the opposite from Benford's Law. I then displayed an image comparing the observed leading digit percentages of the fraudulent charges and Benford's Law. The image, shown in Figure 8(b), reveals a clear difference in the two distributions and shows how payment fraud can be detected. I then stated to the students that currently many companies, countries, and governments are using Benford's Law and statistical inferential techniques to detect fraud.

(4) Shiny App: Benford's Law

I concluded the discussion on Benford's Law by introducing web-based applications I created in Shiny (Chang et al., 2018). Shiny is a web application framework for the programming language R (R Core Team, 2018). With Shiny it is possible to easily create web applications that have similar functionality to Java applets. Doi et al. (2016) provide an introduction on how to use Shiny to create web application teaching tools for statistics. A wide variety of Shiny apps can be found at the "*Shiny App Teaching Tools Collection*" website (https://statistics.calpoly.edu/shiny). Two of the Shiny apps from this site are based on Benford's Law, and I introduced both during the lecture.

The first Shiny app I introduced, titled "Benford's Law: Sequences," examines the leading digit distribution for additive, power, and prime number sequences. The second Shiny app, titled "Benford's Law: Data Examples," examines the leading digit distribution for census data provided by the US Census Bureau and stock data from stock markets around the world (including Japan). A screenshot of the second Shiny app is shown in Figure 9. This figure gives a comparison of observed leading digit percentages for stock prices from the Australia S&P/ASX 200 Stock Market and Benford's Law percentages. The Shiny app also reports a Chi-square goodness-of-fit test to assess the agreement between the two distributions. The large P-value reported in Figure 9 shows that Benford's Law provides a reasonable fit for the distribution of leading digits of Australian stock prices.

4. Topic #3: Longest Run

(1) Motivating Example

For the third lecture topic I began with another activity involving coin tossing (Scheaffer et al. 2004). Students were assigned to one of two groups (Group A and Group

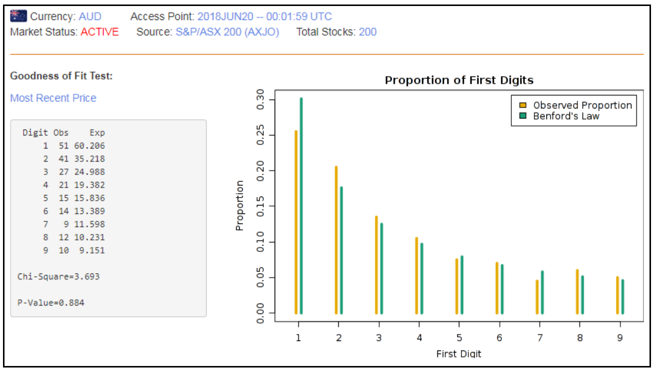


Figure 9. Benford's Law Shiny app screenshot. Display shows the comparison of observed leading digit percentages for stock prices of Australian stocks and Benford's Law percentages.

B). Students in Group A flipped a fair coin 30 times and wrote the outcome for each toss (H = heads, T = tails). Students in Group B were not given a coin and instead wrote the outcomes of what they *imagined* might result from tossing a fair coin 30 times. When the tasks were completed I asked one student from each group to write their results on the board. Figures 10(a) and 10(b) show example outcomes from Group A and Group B respectively. After the outcomes were written on the board I asked all students to compare the results and comment on any differences.

H	H	T	T	T	H	H	H	H	H	T	H	H	T	H
T	H	T	T	H	T	H	H	H	H	T	T	T	T	T

(a)

T	H	H	T	T	H	T	H	H	T	T	H	T	H	H
T	T	T	H	T	H	T	H	T	H	H	H	T	H	H

(b)

Figure 10. (a) Outcomes of 30 actual tosses of a fair coin.
(b) Outcomes of 30 imagined tosses of a fair coin.

Students were quickly able to note that in the outcomes from Group A there are unexpectedly long runs of heads and tails. This is contrasted by the outcomes in Group B which exhibit shorter runs of heads and tails. The short runs in Group B's results indicate that students expected coin toss outcomes to frequently alternate between heads and tails. Note that in Figure 10(a) the maximum run length is 5 whereas in Figure 10(b) the maximum run length is only 3. Was it the case that the unexpected long runs from Group A were due to strange luck or instead that such long runs typically occur when tossing a fair coin? What would we expect to see if we tossed a coin 100 times or even 200 times? To investigate these questions, I then introduced a Shiny app I created to simulate the run length outcomes from a fair coin.

(2) Shiny App: Longest Run

The following Longest Run Shiny app can be found at https://statistics.calpoly.edu/shiny. A screenshot of this app is shown in Figure 11. As shown in the figure, some graphical sliders appear in the left panel of the app. The first slider determines the number of coin tosses to simulate (10 to 400). The second slider controls the minimum run length which determines what particular runs to mark in color. Using this app I showed simulations of 30, 50, 100, and 200 tosses of a fair coin. Also, for each simulation, I showed the occurrence of various run lengths by adjusting the minimum run length slider.

By displaying repeated simulations using this Shiny app the students were able to see that the unexpectedly long runs found in Figure 10(a) are, in fact, not so unusual after all. In addition, the students were also able to note that long runs occur more often than expected. For example, in Figure 11, the outcomes of 200 fair coin tosses are displayed and run lengths of at least 5 are shown in boldface. The app reports that the maximum run length was 8 which, for many students, is unexpectedly high. Also note that there are 11 runs with a run length of at least 5. From this discussion students were able to see that simple random outcomes such as those from fair coins can sometimes behave very unexpectedly.

The Shiny app also provides some statistical inference by reporting the point estimate and a prediction interval for the longest run, as shown at the bottom of Figure 11. Due to the fact that this interval offers a range of reasonable values for the longest run, which is not a parameter, we do not refer to this as a confidence interval but rather as a prediction interval. More details concerning longest run theory can be found in Schilling (1990; 2012).

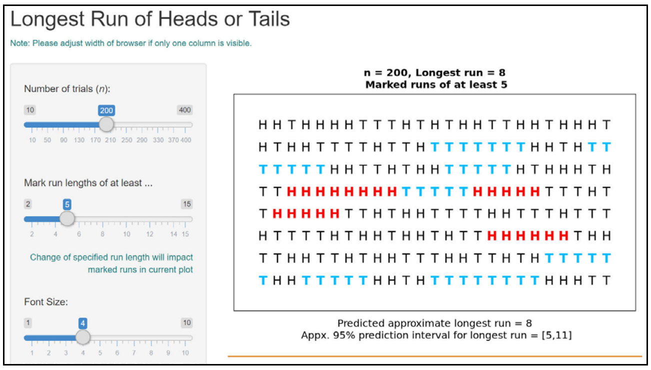
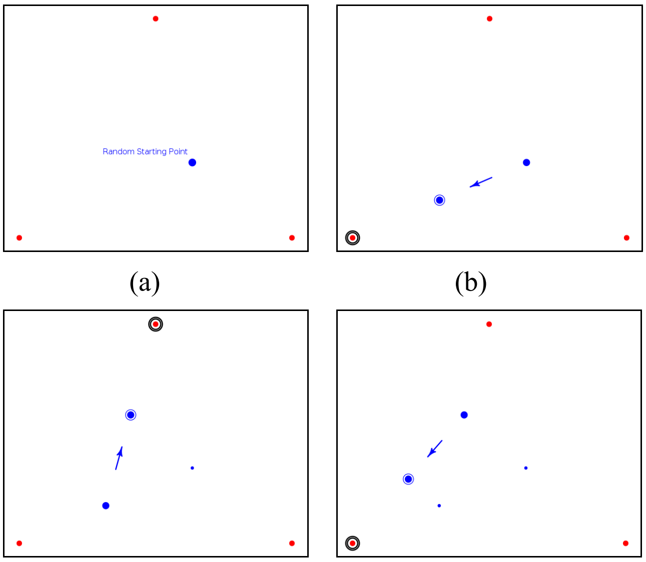


Figure 11. Longest Run Shiny app screenshot. Display shows the outcomes of 200 fair coin tosses where run lengths of at least 5 are shown in boldface. Point estimate and prediction interval of longest run are also reported.



(a) (b)

(c) (d)

Figure 12. (a-d). Progression of the initial steps of the Chaos Game.

5. Topic #4: Chaos Game

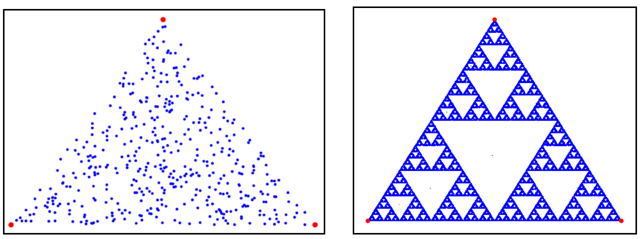
(1) Motivating Example

The final lecture topic was based on what is known as the Chaos Game. This discussion, admittedly, is not related to statistics but it is an interesting example from probability based on a random walk.

I first described how the Chaos Game is played. We start with an image showing the vertices of a triangle and we select a random starting point. It is not required for the starting point to be within the triangle. A triangle vertex is then selected at random. The midway point between the starting point and the selected vertex is marked. A triangle vertex is again selected at random. Then the midway point between the most recently marked point and the selected vertex is marked. One then repeats this process again and again to create many points on the image.

An example of the game is shown in Figure 12. Figure 12(a) shows the triangle vertices and a random starting point. Figure 12(b) shows the next step where the left vertex was randomly selected and the midpoint is marked. Figure 12(c) shows the next step where the upper vertex was randomly selected and the new midpoint is marked. Finally Figure 12(d) shows the next step where the left vertex was randomly selected and the new midpoint is marked.

After explaining the rules of the game, I asked the students to play the game themselves on a sheet of paper and to repeat the process for about 10 steps. Once they had completed this task I asked the students what kind of image they believe would result after 10,000 steps. Most students responded that the final image would be comprised of random points filled uniformly within the triangle and exhibiting no specific pattern. An example of such an image is shown in Figure 13(a). However, the students were shocked to learn that the final image looks like the beautiful and unexpected pattern shown in Figure 13(b). I told the students this is a famous image known as a *fractal* and is specifically known as *Sierpinski's Gasket*. To convince the students of this result I then introduced a Shiny app I created to simulate the Chaos Game.



(a) (b)

Figure 13. (a) The final image most students expect after playing the Chaos Game for many steps. (b) The actual image that results after playing the Chaos Game for many steps.

(2) Shiny App: Chaos Game

The following Chaos Game Shiny app can be found at https://statistics.calpoly.edu/shiny. A screenshot of this app is shown in Figure 14. The default setting of the app is to play the Chaos Game using the vertices of a triangle. In the "Initial Sequence" tab the app allows the user to see the progression of the game for the first 100 steps. In the "Extended Sequence" tab the app allows the user to see the progression from step 100 to step 1,000. In the "Complete Sequence" tab the user can see the progression from step 1,000 to step 50,000. Figure 14 shows the result of the Chaos Game for the first 10,000 steps. As can be seen, the image of Sierpinski's Gasket has appeared.

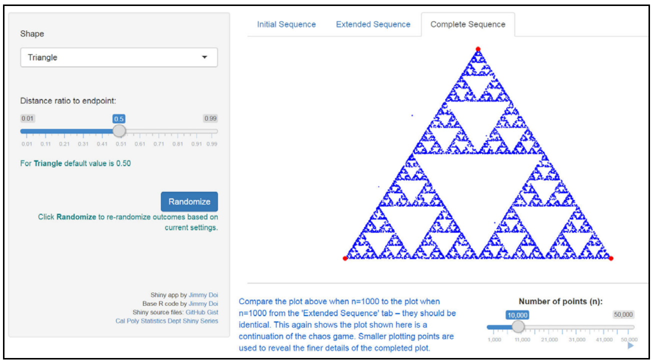


Figure 14. Chaos Game Shiny app screenshot. Display shows the outcomes of the first 10,000 steps of the Chaos Game. The unexpected image is known as Sierpinski's Gasket.

The Shiny app also allows the user to try other variations of the Chaos Game. Instead of the vertices of a triangle, the user can select other shapes such as a square, pentagon, or hexagon. In each case the final result of the Chaos Game yields another form of a beautiful and unexpected fractal.

Finally, I showed the students another Shiny app I created (found at the previously mentioned website) that features the Chaos Game in a three-dimensional setting. In this app the user can simulate the game using the vertices of a tetrahedron, cube, or dodecahedron. In each case the final result yields an unexpected three-dimensional fractal.

6. Lecture Summary

At the end of the lecture I summarized all four lecture topics and how they are all tied together. For each of the lecture topics we discussed, I reminded the students that we encountered an interesting pattern. For the infant experiment example using SBI, we encountered the binomial distribution pattern. For the examples from Benford's Law, we encountered an interesting leading digit pattern. For the longest run example, we encountered an unexpected long run pattern. And for the Chaos Game example, we encountered the surprising fractal-based point location pattern. A common theme that connects all these lecture topics is that, in the presence of randomness, there exist *patterns*. I then mentioned to the students that a major goal of statistics is to understand the patterns induced by randomness so that we may better answer research questions.

Given that this was my first visit to a Super Science High School, I asked the Takasaki High School faculty to describe what their students are like. One of the faculty said, "Our students hope to have careers in medicine, science, and technology. As such, they look to become the leaders of science and technology in Japan." However, the faculty members have also remarked that these students are not always very motivated to study statistics and often ask questions such as "Why do we need to study statistics?" and "How is statistics beneficial for research?" To help emphasize the importance of statistics for the students, I concluded my lecture with the following statement: "Regardless of what you will study, statistics will play a crucial role in your education and your future profession." My hope was that, at least in some small way, through the lecture topics we discussed, I was able to convey to the students how useful and fascinating statistics can be.

7. Conclusion

Lectures using SBI and active learning methods can be very effective to help students better understand statistical and mathematical concepts. Each of the four lecture topics I discussed (SBI, Benford's Law, longest run of heads or tails, Chaos Game) required participation by the students and it seemed everyone was actively involved. Each of the topics had a corresponding web-based application that was easy to use. Given that these applications are freely available by using a web browser, I encouraged all students to continue to use the applications on their own to investigate other application features we were not able to examine during lecture.

Overall, I felt the lecture was successful and well received. Aside from the students, several high school faculty members were also in attendance and I was told that many of the faculty found the lecture to be informative and interesting. In addition, their faculty mentioned they were planning to adopt my lecture presentation materials of all four topics for their math courses. Recently I have received other invitations to give this presentation and the next presentation will be during summer 2019 at Kanonji Super Science High School in Kagawa, Japan.

Acknowledgments

I would like to thank Dr. Michiko Watanabe and Mr. Yasuhiko Nakajima for their many efforts to make the lecture at Takasaki High School possible. I also thank Drs. Ann Watkins, Roxy Peck, Beth Chance, and Mark Schilling on valuable suggestions provided on earlier versions of this article. Finally, I am indebted to Dr. Michiko Watanabe and Dr. Keiichi Nishimura for their invaluable Japanese translation assistance for this article.

REFERENCES

American Statistical Association. (2007). Guidelines for Assessment and Instruction in Statistics Education (GAISE): Pre-K-12 Report. Alexandria, VA: American Statistical Association. https://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf.

American Statistical Association. (2016). Guidelines for Assessment and Instruction in Statistics Education (GAISE): College report. Alexandria, VA: American Statistical Association. https://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf.

Chance, B, Ben-Zvi, D, Garfield, J, and Medina, E. (2007) "The Role of Technology in Improving Student Learning of Statistics," Technology Innovations in Statistics Education, 1(1).

Chance B, and Rossman A. (2018). *Investigating Statistical Concepts, Applications, and Methods* (3rd Ed), Publisher: Author.

Chance, B, and McGaughey, K. (2014). "Impact of a Simulation/Randomization-Based Curriculum on Student Understanding of p-Values and Confidence Intervals," Proceedings of the 9th International Conference on Teaching Statistics. https://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_6B1_CHANCE.pdf.

Chang, W, Cheng, J, Allaire, JJ, Xie, Y and McPherson, J. (2018). shiny: Web Application Framework for R. R package version 1.2.0. https://CRAN.R-project.org/package=shiny.

delMas, R, Garfield, J, and Chance, B. (1999), "A Model of Classroom Research in Action: Developing Simulation Activities to Improve Students' Statistical Reasoning," Journal of Statistics Education, 7(3).

Diez, D, Barr, C, and Çetinkaya Rundel, M. (2014), *Introductory Statistics With Randomization and Simulation* (1st Ed), CreateSpace Independent Publishing Platform. https://www.openintro.org/stat/textbook.php?stat_book=isrs.

Doi, J, Potter, G, Wong, J, Alcaraz, A, and Chi, P. (2016). "Web Application Teaching Tools for Statistics Using R and Shiny". Technology Innovations in Statistics Education, 9(1).

Franklin, C and Garfield, J. (2006). The GAISE project: Developing statistics education guidelines for pre K-12 and college courses. In G. Burrill (Ed.), Thinking and reasoning with data and chance: 2006 NCTM yearbook (p. 435–475). Reston, VA: National Council of Teachers of Mathematics.

Garfield, J, and Ben-Zvi, D. (2008), *Developing*

Students' Statistical Reasoning: Connecting Research and Teaching Practice, Kluwer Academic Publishers.

Hamlin, J, Wynn, K, and Bloom, P. (2007) "Social evaluation by preverbal infants" Nature 450, 557–9.

Hildreth L, Robison-Cox J, and Schmidt J. (2018). "Comparing Student Success and Understanding in Introductory Statistics under Consensus and Simulation-Based Curricula". Statistics Education Research Journal, 17(1), 103-120. https://iase-web.org/documents/SERJ/SERJ17(1)_Hildreth.pdf.

Lock, R, Lock, P, Morgan, K, Lock, E, and Lock, D. (2017). *Statistics: Unlocking the Power of Data* (2nd Ed), Wiley.

Luque, B and Lacasa, L. (2009) "The first-digit frequencies of prime numbers and Riemann zeta zeros". Proc. R. Soc. A , 465, 2197–2216.

Nigrini, M. (1999). "I've Got Your Number". Journal of Accountancy, 187(5), 79–83. http://www.journalofaccountancy.com/issues/1999/may/nigrini.html.

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roy S, Rossman A, Chance B, Cobb G, VanderStoep J, Tintle N and Swanson T. (2014). "Using simulation/randomization to introduce p-value in week 1." Proceedings of the ninth International Conference on Teaching Statistics. http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_4A2_ROY.pdf.

Scheaffer R, Watkins A, Witmer J, and Gnanadesikan M. (2004) *Activity-Based Statistics* (2nd Ed), Key College

Schilling, M. (1990). "The Longest Run of Heads," The College Mathematics Journal, 21(3), 196–207.

Schilling, M. (2012). "The Surprising Predictability of Long Runs," Mathematics Magazine, 85(2), 141–149.

Tintle N, Carver R, Chance B, Cobb G, Rossman A, Roy S, Swanson T, and VanderStoep J. (2018) *Introduction to Statistical Investigations AP Edition*, John Wiley & Sons.

Tintle N, Chance B, Cobb G, Rossman A, Roy S, Swanson T, and VanderStoep J. (2016) *Introduction to Statistical Inference*, John Wiley & Sons.

Tintle N, Chance B, Cobb G, Roy S, Swanson T, and VanderStoep J. (2015). "Combating antistatistical thinking using simulation-based methods throughout the undergraduate curriculum" The American Statistician, 69(4):362–70.

Tintle, N, Rogers, A, Chance, B, Cobb, G, Rossman, A, Roy, S, … VanderStoep, J. (2014). "Quantitative Evidence for the Use of Simulation and Randomization in the Introductory Statistics Course," Proceedings of the 9th International Conference on Teaching Statistics.

Tintle, N, VanderStoep, J, Holmes, V, Quisenberry, B, and Swanson, T. (2011). "Development and Assessment of a Preliminary Randomization-Based Introductory Statistics Curriculum," Journal of Statistics Education, 19(1).