Simulation-Based Inference Problems

(Problems that could be solved using simulation-based methods)

1. Can Chimps Solve Problems?

In 1978, researchers Premack and Woodruff[1] reported a study in *Science* magazine where an adult chimpanzee named Sarah was shown videotapes of eight different scenarios of a human being faced with a problem. After each videotape showing, she was presented with two photographs, one of which depicted a correct solution to the problem. For example, if the videotape showed the human unable to extricate himself from a locked cage, the photograph with the correct solution showed a key.

Sarah picked the photograph with the correct solution for seven of the eight problems. ***Can chimpanzees solve problems?***

[1] **"Chimpanzee problem-solving: A test for comprehension,"** ***Science*** **3, vol 202(4367), pp. 532-535**

2. Can Dolphins Communicate?

(From ISI Workshop[2] at Butler University on November 1, 2014)

Phase 1: Dr. Jarvis Bastian placed a large canvas curtain in the middle of the pool. Doris was on one side of the curtain and could see the headlight, whereas Buzz was on the other side of the curtain and could not see the headlight. Dr. Bastian turned on the headlight and let it shine steadily. He then watched to see what Doris would do. After looking at the light, Doris swam near the curtain and began to whistle loudly. Shortly after that, Buzz whistled back and then pressed the button on the right—he got it correct and so both dolphins got a fish. But this single attempt was not enough to convince Dr. Bastian that Doris had communicated with Buzz through her whistling. Dr. Bastian repeated the process several times, sometimes having the light blink (so Doris needed to let Buzz know to push the left button) and other times having it glow steadily (so Doris needed to let Buzz know to push the right button). He kept track of how often Buzz pushed the correct button.

Depending whether or not the light was blinking or shown steadily, Doris had to communicate to Buzz as to which button to push. If he pushed the right button, they both got food (fish).

In one phase of the study, Dr. Bastian had Buzz attempt to push the correct button a total of 16 different times. In this phase of the study, Buzz pushed the correct button 15 out of 16 times. Does this provide convincing evidence the Doris and Buzz are communicating?

Phase 2: The study design is similar with some adjustments to the barrier between Doris and Buzz, with the canvas curtain being replaced by a plywood board. The research conjecture, observational units, and variable remain the same.

In this case, Buzz pushed the correct button only 16 out of 28 times. The variable is the same (whether or not Buzz pushed the correct button), but the number of observational units has changed to 28 (the number of attempts). Does this provide convincing evidence the Doris and Buzz are communicating?[3]

---

[2] www.rossmanchance.com/isi/DolphinsHS.docx

[3] (Epilogue: Dr. Bastian soon discovered that in the previous set of attempts, the equipment malfunctioned and the food dispenser for Doris did not operate and so Doris was not receiving her fish rewards during the study. Dr. Bastian fixed the equipment and ran the study again. This time he found convincing evidence that Buzz was not guessing.)

3. Left Twix or Right Twix?

Mr. Ferris's AP stats students found that 66% of randomly surveyed NHS students (n=50) preferred Left Twix over Right Twix[4]. Is this a significant difference?

[4] Many students articulated specific and clear reasons for their choice based on chocolate, caramel, taste, etc.

4. The Kristen Gilbert Case

(www.stat.ucla.edu/~nchristo/statistics100B/article.pdf)

Kristen Gilbert worked as a nurse in the intensive care unit of the VA hospital in Northampton, Massachusetts in the 1990's. During her shifts, other nurses became suspicious that she was killing patients by injecting them with epinephrine, a heart stimulant. An analysis of 1641 eight-hour shifts was presented as evidence at her trial (Cobb and Gelbach, 2005). Is there statistical evidence that Gilbert murdered patients?

	Death on shift?		
Gilbert Present?	Yes	No	Total
Yes	40	217	257
No	34	1350	1384
Total:	74	1567	1641

5. Hiring Discrimination?

(from TPS5e, page 5)

An airline has just finished training 25 pilots— 15 male and 10 female—to become captains. Unfortunately, only eight captain positions are available right now. Airline managers announce that they will use a lottery to determine which pilots will fill the available positions. The names of all 25 pilots will be written on identical slips of paper. The slips will be placed in a hat, mixed thoroughly, and drawn out one at a time until all eight captains have been identified.

A day later, managers announce the results of the lottery. Of the 8 captains chosen, 5 are female and 3 are male. Some of the male pilots who weren't selected suspect that the lottery was not carried out fairly. One of these pilots asks your statistics class for advice about whether to file a grievance with the pilots' union.

The key question in this possible discrimination case seems to be: *Is it plausible* (*believable*) *that these results happened just by chance?*

6. The Midge Problem

(From Dan Teague, NCSSM)

In 1981, two new varieties of a tiny biting insect called a midge were discovered in the jungles of Brazil by biologists W. L. Grogan and W. W. Wirth. They dubbed one kind of midge an *Apf* midge and the other an *Af* midge. The biologists found that the *Apf* midge is a carrier of a debilitating disease that causes swelling of the brain when a human is bitten by an infected midge. Although the disease is rarely fatal, the disability caused by the swelling may be permanent. The other form of the midge, the *Af*, is quite harmless and a valuable pollinator. In an effort to distinguish the two varieties, the biologists took measurements on the midges they caught. The two measurements taken were wing length and antenna length, both measured in centimeters. The data are provided below.

***Af* midges**

Wing length (cm)	1.72	1.64	1.74	1.7	1.82	1.82	1.9	1.82	2.08
Antenna length (cm)	1.24	1.38	1.36	1.4	1.38	1.48	1.38	1.54	1.56

***Apf* midges**

Wing length (cm)	1.78	1.86	1.96	2.0	2.0	1.96
Antenna length (cm)	1.14	1.2	1.3	1.26	1.28	1.18

Is it possible to distinguish an *Af* midge from an *Apf* midge on the basis of wing and antenna lengths?[5]

[5] **Source: "The Midge Problem," *Everybody's Problems, Consortium,* Number 55, Fall, 1995, COMAP, Inc., Lexington, MA.**

7. Dietary Change and Cholesterol

From NCTM's *Navigating through Data Analysis in Grades 9-12*

High cholesterol is a contributor to heart disease. Table 1 lists data from a study investigating the effect of dietary change on cholesterol levels. Twenty-four hospital employees voluntarily switched from "a standard American diet" to a vegetarian diet for one month. The data show their cholesterol levels both before and after the dietary change, in milligrams of cholesterol per deciliter of blood (mg/dL). Suppose for the activities that it is always desirable to decrease the level of cholesterol in the blood. Thus, assume that the purpose of the switch to the new vegetarian diet was to decrease that level.

Table 1

Cholesterol Levels before and after Changing Diets

Before (mg/dL)	After (mg/dL)	Before (mg/dL)	After (mg/dL)
195	146	169	182
145	155	158	127
205	178	151	149
159	146	197	178
244	208	180	161
166	147	222	187
250	202	168	176
236	215	168	145
192	184	167	154
224	208	161	153
238	206	178	137
197	169	137	125

(From Rosner, Bernard, *Fundamentals of Biostatistics* [Boston: Duxbury Press, 1986])

Was the diet a success, or could decreases in cholesterol levels have been due merely to chance?

1. Write 2-3 questions that could be answered *from this data set.*

2. Write 2-3 questions that this data set could answer if someone asked YOU to change your diet to lower your cholesterol.

3. Answer one of the questions you wrote using graphs and statistics from the data. Explain and justify your conclusion.

4. To whom can conclusions from this data be generalized (applied)? Explain.

8. Stent Study

(from OpenIntro v.3)

Stents are devices put inside blood vessels that assist in patient recovery after cardiac events and reduce the risk of an additional heart attack or death. Many doctors have hoped that there would be similar benefits for patients at risk of stroke.[6] We start by writing the principal question the researchers hope to answer:

Does the use of stents reduce the risk of stroke?

The researchers who asked this question collected data on 451 at-risk patients. Each volunteer patient was randomly assigned to one of two groups:

Treatment group: Patients in the treatment group received a stent and medical management. The medical management included medications, management of risk factors, and help in lifestyle modification.

Control group: Patients in the control group received the same medical management as the treatment group, but they did not receive stents.

Researchers randomly assigned 224 patients to the treatment group and 227 to the control group. In this study, the control group provides a reference point against which we can measure the medical impact of stents in the treatment group.

Researchers studied the effect of stents at two time points: 30 days after enrollment and 365 days after enrollment. The results of 5 patients and the overall summary are in the tables below.

Patient	group	0-30 days	0-365 days
1	treatment	no event	no event
2	treatment	stroke	stroke
3	treatment	no event	no event
⋮	⋮	⋮	
450	control	no event	no event
451	control	no event	no event

	0-30 days		0-365 days	
	stroke	no event	stroke	no event
treatment	33	191	45	179
control	13	214	28	199
Total	46	405	73	378

Table 1.2: Descriptive statistics for the stent study.

[6] Chimowitz MI, Lynn MJ, Derdeyn CP, et al. 2011. Stenting versus Aggressive Medical Therapy for Intracranial Arterial Stenosis. New England Journal of Medicine 365:993-1003. www.nejm.org/doi/full/10.1056/NEJMoa1105335. NY Times article reporting on the study: www.nytimes.com/2011/09/08/health/research/08stent.html.

9. Is Yawning Contagious?

(TPS 4e)

Data from an experiment involving 50 subjects on the TV show *MythBusters* investigated this question. Here's a brief recap. Each subject was placed in a booth for an extended period of time and monitored by hidden camera. Thirty-four subjects were given a "yawn seed" by one of the experimenters; that is, the experimenter yawned in the subject's presence before leaving the room. The remaining 16 subjects were given no yawn seed. What happened? The table below shows the results:

	Yawn Seed?		
Subject Yawned?	**Yes**	**No**	**Total**
Yes	10	4	14
No	24	12	36
Total	34	16	50

So…is yawning contagious? Explain.

10. Can Facebook Identify Terrorists?[7]

Suppose Facebook has developed an algorithm that can identify Facebook users who are twice as likely to be involved in terrorist activities as the typical Facebook user. Your neighbor has been identified as one of the 100,000 Facebook users that fit the profile of a terrorist. Assuming there are 200 million Facebook users, 10,000 of which really are terrorists, should you be worried about your neighbor? Explain.[8]

[7] Ellenberg, Jordan (2014-05-29). How Not to Be Wrong: The Power of Mathematical Thinking (p. 167). Penguin Group US. Kindle Edition.

[8] Two questions worth considering: 1) What is the likelihood that your neighbor would make the list if, in fact, he is not a terrorist? 2) What is the likelihood that your neighbor is not a terrorist given that he is on the list?

11. Halloween Treats

(From ISI Workshop at Butler University on November 1, 2014)

Stemming from concern over the nation's obesity epidemic, researchers investigated whether children might be as tempted by toys as by candy for Halloween treats. Test households in five Connecticut neighborhoods offered children two plates: one with lollipops or fruit candy and one containing small, inexpensive Halloween toys, like plastic bugs that glow in the dark. The researchers observed the selections of 283 trick-or-treaters between the ages of 3 and 14 (Schwartz, Chen, and Brownell, 2003). The researchers found that 148 trick-or-treaters selected the candy, and 135 chose a toy.

Calculate the p-value using simulation and theory-based methods.

12. Colon Cancer

(2015 Exam #4)

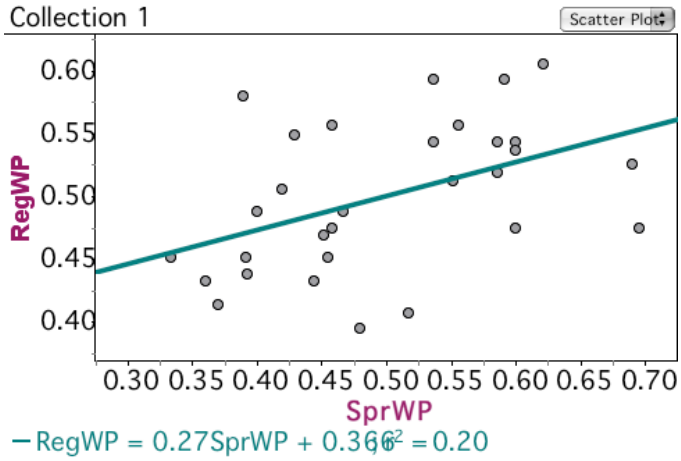
A researcher conducted a medical study to investigate whether taking a low-dose aspirin reduces the chance of developing colon cancer. As part of the study, 1,000 adult volunteers were randomly assigned to one of two groups. Half of the volunteers were assigned to the experimental group that took a low-dose aspirin each day, and the other half were assigned to the control group that took a placebo each day. At the end of six years, 15 of the people who took the low-dose aspirin had developed colon cancer and 26 of the people who took the placebo had developed colon cancer. At the significance level a 0.05, do the data provide convincing statistical evidence that taking a low-dose aspirin each day would reduce the chance of developing colon cancer among all people similar to the volunteers?

13. MLB Spring Training vs. Regular Season Games

(From Josh Tabor at ISI Workshop at Butler University on November 1, 2014)

Like many sports, Major League Baseball teams play spring training exhibition games before the regular season begins. These games do not count in the regular season standings, but some fans believe that a team's performance in spring training games can predict their team's success (or lack thereof) in the regular season. Others believe that spring training games cannot predict the team's performance in the regular season.

The data provided show the 2014 winning proportions for all 30 MLB teams during both spring training and the regular season.



Team	SprWP	RegWP
Arizona	0.480	0.395
Atlanta	0.400	0.488
Baltimore	0.591	0.593
Boston	0.393	0.438
Chicago Cubs	0.455	0.451
Chicago Sox	0.391	0.451
Cincinnati	0.452	0.469
Cleveland	0.690	0.525
Colorado	0.517	0.407
Detroit	0.556	0.556
Houston	0.444	0.432
Kansas City	0.429	0.549
LA Angels	0.621	0.605
LA Dodgers	0.389	0.580
Miami	0.600	0.475
Milwaukee	0.419	0.506
Minnesota	0.360	0.432
NY Mets	0.467	0.488
NY Yankees	0.586	0.519
Oakland	0.536	0.543
Philadelphia	0.333	0.451
Pittsburgh	0.600	0.543
San Diego	0.458	0.475
San Francisco	0.586	0.543
Seattle	0.60	0.537
St. Louis	0.458	0.556
Tampa Bay	0.696	0.475
Texas	0.370	0.414
Toronto	0.552	0.512
Washington	0.536	0.593

1. Does there appear to be a positive association between spring and regular games? Explain.

2. There are generally two explanations for this positive association. What are they?

3. One of these explanations *might* be eliminated from consideration by using a randomization-based approach. Explain.

14. Fun Size M&M's Bags

Collect a "class set" of the counts of M&M's in fun size bags. Graph and find the mean. What would be a plausible range for the mean number of candies in a fun size bag? How could we determine this range? Perhaps by collecting more "class sets" of counts (more chocolate!). OR, we could conduct a simulation with the ONE sample we have already collected (cheaper and easier). Use StatKey (online tool) to create a bootstrap confidence interval for the mean number of M&M's in a fun size bag.