S-2016

Section 5.1 - Randomness, Probability & Simulation (pp. 287-299)

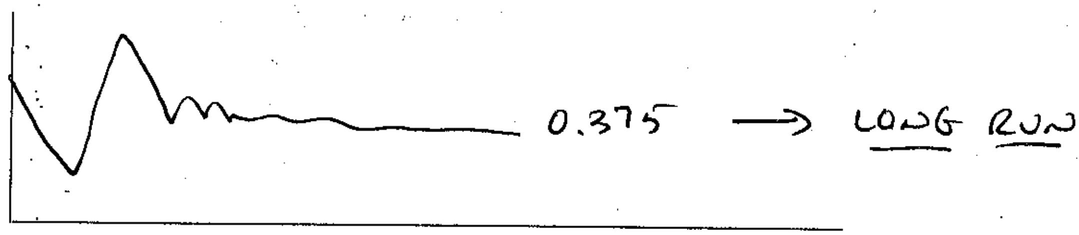
1. The Idea of Probability

- Random samples and randomized experiments → REVIEW
- Avoid bias by allowing chance to decide what individuals get selected.
- Chance behavior is *unpredictable in the* ***short run*** *but has a regular and predictable pattern in the* ***long run.*** * LONG RUN BEHAVIOR!
- This is the basis for **probability.**

* AP STATS IS NOT A PROBABILITY COURSE – DISCUSS.

Application - Random Babies

Suppose a stork randomly delivers four babies to four different houses. We are going to *simulate* this situation and count the number of correct deliveries.



(www.rossmanchance.com/applets/randomBabies/Babies.html)

Plot the proportion of proportion of trials where there were 0 matches. What are we observing when we run the simulation a large number of times? "CONVERGENCE"

ABOUT 35% OF THE TIME NONE OF THE BABIES WOULD END UP @ CORRECT HOUSE.

Law of Large Numbers - The fact that the proportion of trials that had no babies delivered to the right house converges to 0.375 is guaranteed by the **law of large numbers.** This result says that if we observe more and more repetitions of a chance process, the proportion of times that a specific outcome occurs approaches a single value. The single value is called **probability**.

What implications does this have on sampling design and experimental design? – DISCUSS

* LARGE SAMPLES ✓
* REPLICATION ✓

Definition: The **probability** of any outcome of a chance process is a number between 0 and 1 that describes the proportion of times the outcome would occur in a very long series of repetitions.

"When it all comes down we will still come through -- In the long run." *Eagles, "In the Long Run"*

$0 \leq P(x) \leq 1$ **

Example. How much should a company charge for an extended warranty for a specific type of cell phone? Suppose that 5% of these cell phones under warranty will be returned, and the cost to replace the phones is $150. If the company knew which phones would go bad, it could charge $150 for these phones and $0 for the rest. However, since the company cannot know which phones will be returned but knows that about 1 in every 20 will be returned. How much should they charge for the extended warranty?

THEY SHOULD CHARGE AT LEAST $150/20 = $7.50 FOR EXTENDED WARRANTY.

Other examples:

INSURANCE COVERAGE.

CHECK YOUR UNDERSTANDING

1. According to the "Book of Odds," the probability that a randomly selected U.S. adult usually eats breakfast is 0.61.

(a) Explain what probability 0.61 means in this setting. 61% ANSWER "YES"

(b) Why doesn't this probability say that if 100 U.S. adults are chosen at random, exactly 61 of them usually eat breakfast? → # VARIES FROM SAMPLE TO SAMPLE.

2. Probability is a measure of how likely an outcome is to occur. Match one of the probabilities that follow with each statement. Be prepared to defend your answer.

0 0.01 0.3 0.6 0.99 1

(a) This outcome is impossible. It can never occur. 0

(b) This outcome is certain. It will occur on every trial. 1

(c) This outcome is very unlikely, but it will occur once in a while in a long sequence of trials. 0.01

(d) This outcome will occur more often than not. 0.6, 0.99

---

2. Myths about Randomness - The idea of probability seems straightforward. It answers the question "What would happen if we did this many times?" In fact, both the behavior of random phenomena and the idea of probability are a bit subtle. We meet chance behavior constantly, and psychologists tell us we deal with it poorly.

Application - Suppose that a basketball announcer suggests that a certain player is *streaky*. That is, the announcer believes that if the player makes a shot, then he is more likely to make his next shot. As evidence, he points to a recent game where the player took 30 shots and had a streak of 7 in a row. Is this evidence of streakiness or could it have occurred by chance? Assuming the player makes 50% of his shots and the results of a shot do not depend on previous shots, how likely is it for the player to have a streak of 7 or more made shots in a row?

Solution (4 Step Process):

1. **State** - How likely is it for the player to have a streak of 7 or more made shots in a row?

2. **Plan** - Use the random number generator on the calculator to generate 30 random 0s and 1s. 0 = misses shot, 1 = makes shot. Record the outcome of each trial. Record the longest streaks on a dot plot. RANDINT (0, 1, 30)

3. **Do** - Have each student do this 2 times.

TYPO

0 1 2 3 4 5 6 7 8 9 10 11 12

Longest Streak

4. **Conclude**

---

3. **Simulations** - The application that we just conducted is called a **simulation**. In fact, the vaunted Hyena Problem on day 1 was a simulation. To perform a simulation, we are going to use the venerable 4 Step Process.

Performing a Simulation

1. **State** - What is the question of interest about some chance process?

2. **Plan** - Describe how to use a *chance device* to imitate one repetition of the process. Explain clearly how to identify the outcomes of the chance process and what variable to measure.

3. **Do** - Perform *many* repetitions of the simulation. (At least 30.)

4. **Conclude** - Use the results of the simulation to answer the question of interest.

IN CONTEXT !!

Examples of 4-Step Process

Refer to the table on p. 296

Examples on pp. 296-297

Application - At a department picnic, 18 students in the mathematics/statistics department at a university decide to play a softball game. Twelve of the 18 students are math majors and 6 are stats majors. To divide into two teams of 9, one of the professors put all the players' names into a hat and drew out 9 players to form one team, with the remaining 9 players forming the other team. The players were surprised when one team was made up entirely of math majors. Is it possible that the names were not adequately mixed in the hat, or could this have happened by chance? Design and carry out a simulation to help answer this question.

1. STATE - WHAT IS PROB. THAT WHEN RANDOMLY ASSIGNING 12 MATH MAJORS + 6 STATS MAJORS TO 2 TEAMS, THERE WILL BE ONE TEAM W/ALL MATH MAJORS?

2. PLAN - USING A RANDOM # GENERATOR, GENERATE RANDOM INTEGERS 1-18, 1-12 = MATH, 13-18 = STATS; 1st 9 #'S GO TO 1st TEAM, REMAINDER GO TO 2ND TEAM. COUNT # OF MATH MAJORS ON EACH TEAM + RECORD RESULTS. PLOT RESULTS IN A DOT PLOT.

3. DO: (2 TRIALS PER STUDENT)

   RESULTS:

4. CONCLUDE:

HW - 1, 3, 7, 9, 11, 15, 17, 23, 25, 38*