EXPLORATIONS ™

Activity 7

Objectives

- Explain the relationship between relative frequency and theoretical probability
- Confront common misconceptions about probabilities of streaks

Gambler's Fallacy: Longest Streaks

Introduction

Do you think that a fair coin would ever land heads up 27 times in a row? The FlipSite Web site (http://www.mcli.dist.maricopa.edu/mobius/flip/) keeps a database of the results of a coin flip that occurs every 5 seconds from January 1 through December 31 of each year. Explore this site to see how often a streak of 27 heads occurred for this year.

In this activity you will explore longest streaks for seven tosses of a coin. For example, HTHTHHH has three heads in a row for its longest streak. HTTTTTH has a streak of five tails.

Problem

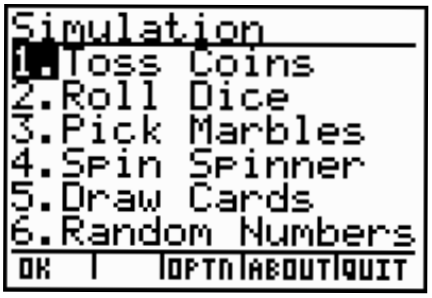
Do you think that having a longest streak of four is more or less likely than a streak of three? Is a longest streak of three more or less likely than a streak of two? How likely is a streak of six out of seven tosses? Will a streak of seven out of seven tosses ever occur?

Exploration

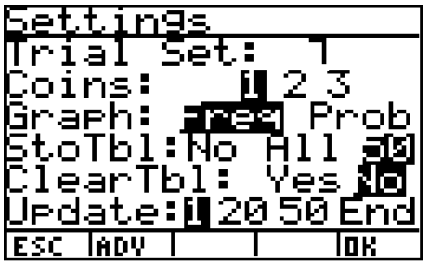
Simulate 30 Sets of 7 Coin Tosses

Respond to # 1 – 4 on the Student Worksheet.

1. Make sure that the program STREAKS has been loaded onto your graphing calculator. Your teacher will give you instructions on how to load programs onto your graphing calculator.
2. Open the Probability Simulation App and select **Toss Coins.**



3. Select **SET** and change the settings as shown.



4. Select **OK** to return to the Simulation screen.
5. Select **TABL** so that the data is shown in a table.
6. Select **TOSS** to run the simulation.

   Count the longest uninterrupted streak (either heads or tails). Place a tally mark in the corresponding column of # 5 on the Student Worksheet. If there are two streaks of the same length within a set, count only one outcome. For example, for the following outcome, THHHTTT, place one tally mark in the column labeled 3.
7. Record the results for 30 trials. To toss a new set of seven coins, open the **Settings** menu and set **ClearTbl** to **Yes**.

Respond to # 5 – 6 on the Student Worksheet.

Simulating 500 Sets of Seven Tosses

Perhaps you are surprised by some of the results. Since 30 trials are not very many, your results may be due to experimental variance. The program STREAKS7 will collect data from 500 sets of seven tosses.

Run the STREAKS7 program and record the results in the chart of # 7 on the Student Worksheet.

As the program runs, it displays the set number currently being tallied. The program stores the number of times each streak occurred from 500 sets of seven tosses in L1. For example, the value in L1 (3) represents the number of times a streak of 3 was the longest streak.

Respond to # 7 on the Student Worksheet.

Combining Many Sets of Seven Tosses

Pool the class results of running the STREAKS7 program.

Respond to # 8 – 9 on the Student Worksheet.

Finding the Theoretical Probabilities

Now that you have run simulations to get experimental results, you will want to see how closely the experimental probabilities match the theoretical probabilities. The 2^7 or 128 outcomes for seven tosses of a coin are organized in a chart on the next page.

Use the Sample Space for 7 Coin Tosses chart on the next page to record the length of the longest uninterrupted run for each combination.

Respond to # 10 – 12 on the Student Worksheet.

Extension

Explore longest streaks of choosing (with replacement) the same color marble in seven picks from a bag of three different colored marbles . Before beginning this exploration, make conjectures using the results from the coin toss exploration. For example, the probability for a streak of three was slightly greater than a steak of two but quite a bit greater than that for a streak of four. Do you think the marble simulation will follow a similar pattern?

You may choose to explore this problem theoretically or experimentally. You may want to modify the STREAKS7 program to use for this scenario.

Sample Space for 7 Coin Tosses

H = Heads O = Tails

HHHHHHH	HHHHHOH	HHHHHHO	HHHHHOO
OHHHHHH	OHHHHOH	OHHHHHO	OHHHHOO
HOHHHHH	HOHHHOH	HOHHHHO	HOHHHOO
OOHHHHH	OOHHHOH	OOHHHHO	OOHHHOO
HHOHHHH	HHOHHOH	HHOHHHO	HHOHHOO
OHOHHHH	OHOHHOH	OHOHHHO	OHOHHOO
HOOHHHH	HOOHHOH	HOOHHHO	HOOHHOO
OOOHHHH	OOOHHOH	OOOHHHO	OOOHHOO
HHHOHHH	HHHOHOH	HHHOHHO	HHHOHOO
OHHOHHH	OHHOHOH	OHHOHHO	OHHOHOO
HOHOHHH	HOHOHOH	HOHOHHO	HOHOHOO
OOHOHHH	OOHOHOH	OOHOHHO	OOHOHOO
HHOOHHH	HHOOHOH	HHOOHHO	HHOOHOO
OHOOHHH	OHOOHOH	OHOOHHO	OHOOHOO
HOOOHHH	HOOOHOH	HOOOHHO	HOOOHOO
OOOOHHH	OOOOHOH	OOOOHHO	OOOOHOO
HHHHOHH	HHHHOOH	HHHHOHO	HHHHOOO
OHHHOHH	OHHHOOH	OHHHOHO	OHHHOOO
HOHHOHH	HOHHOOH	HOHHOHO	HOHHOOO
OOHHOHH	OOHHOOH	OOHHOHO	OOHHOOO
HHOHOHH	HHOHOOH	HHOHOHO	HHOHOOO
OHOHOHH	OHOHOOH	OHOHOHO	OHOHOOO
HOOHOHH	HOOHOOH	HOOHOHO	HOOHOOO
OOOHOHH	OOOHOOH	OOOHOHO	OOOHOOO
HHHOOHH	HHHOOOH	HHHOOHO	HHHOOOO
OHHOOHH	OHHOOOH	OHHOOHO	OHHOOOO
HOHOOHH	HOHOOOH	HOHOOHO	HOHOOOO
OOHOOHH	OOHOOOH	OOHOOHO	OOHOOOO
HHOOOHH	HHOOOOH	HHOOOHO	HHOOOOO
OHOOOHH	OHOOOOH	OHOOOHO	OHOOOOO
HOOOOHH	HOOOOOH	HOOOOHO	HOOOOOO
OOOOOHH	OOOOOOH	OOOOOHO	OOOOOOO

STREAKS7 Program

STREAKS7 is a program written for the TI-83 Plus graphing calculator that will tabulate the longest streaks occurring in 500 sets of seven tosses of a coin. List $L_1(1)$ stores how many sets had a longest streak of 1; List $L_1(2)$, a longest streak of 2, and so forth. List L_2 is used to store a single set of seven tosses. The calculator screen displays the set number that has just been tossed. When the program is done executing, press [STAT] and select **1:Edit** to view the results in L_1.

If you want to check to make sure the program is tabulating the longest length correctly, insert a quote (") at the beginning of Lines 3, 4 and 20. This will block out the loop of 500 sets; thus, the program will only execute one set of seven tosses. Run the program several times, checking L_1 after each execution. You should find a 1 in L_1 in the cell that represents the longest streak length found in L_2. When you are convinced the program is doing what it is meant to do, delete the quotes you inserted.

The toughest part of the code to understand is the section that determines the longest streak length in a set of 7 tosses, Lines 10-18. However, it is merely doing what you would do if you could only see one toss at a time instead of all seven at once.

LINE	PROGRAM CODE	KEYSTROKE	ACTION
Line 01	`ClrList L1,L,`	[STAT] 4:	Clears the data from L_1 and L_2
Line 02	`7→dim(L1)`	[2nd] [LIST] ▶ OPS 3:	Fills the first 7 cells of L_1 with 0's
Line 03	`For(J,1,500)`	[PRGM] 4:	Starts the loop of 500 trials of 7 tosses
Line 04	`Disp J`	[PRGM] ▶ I/O 3:	Displays the set number, J, that will be tossed
Line 05	`For(I,1,7)`		Starts the loop of 7 tosses in a trial
Line 06	`randInt(0,1)→L,(I)`	[MATH] ▶ ▶ ▶ PRB 5:	Stores the coin toss in L_2; 0 for tails, 1 for heads
Line 07	`End`	[PRGM] 7:	Ends the loop for a trial of 7 tosses
Line 08	`1→N`	[STO▶]	N keeps track of the length of a new streak in the set
Line 09	`1→S`		S keeps track of the length of the longest streak
Line 10	`For(I,1,6)`		Starts a loop comparing one toss to the next
Line 11	`If L,(I)=L,(I+1)`	[PRGM] 1:	Checks if the one toss in the trial is equal to the next
Line 12	`Then`	[PRGM] 2:	If they are equal then
Line 13	`N+1→N`		1 is added to the length counter
Line 14	`Else`	[PRGM] 3:	Else if they are not equal
Line 15	`1→N`		The length counter is set to 1 again
Line 16	`End`		Ends the if-then-else condition

Line 17	`max(N,S)→S`	MATH ▸ NUM 7:	Sets the longest streak equal to the max of N and S
Line 18	`End`		Ends the loop comparing one toss to the next
Line 19	`L1(S)+1→L1(S)`		Adds 1 to the appropriate cell in L for the longest streak in a trial of 7 tosses
Line 20	`End`		Ends the loop for 500 trials

Student Worksheet

Name _____

Date _____

Gambler's Fallacy: Longest Streaks

Respond to # 1 – 4 based on your intuition. Do not make any calculations.

1. Which is more likely?

a. A hand of cards with 4 kings and an ace

b. A hand of cards with an ace, 2, and 8 of clubs, and 10 and jack of diamonds

2. In a set of seven tosses of a fair coin, which is more likely?

a. a sequence of HHHHHHH

b. a sequence of HTHTHTH

3. In seven tosses of a fair coin, which is more likely?

a. a longest heads streak of one

b. a longest streak of one

4. A coin is tossed multiple sets of seven times and the longest uninterrupted streak of each set is recorded. Without doing any calculations, what probability (to the nearest percent) would you predict for each outcome?

STREAK of:	1	2	3	4	5	6	7	Total
Prediction								

5. Run the simulation of seven rolls 30 times. After each set of seven rolls, put a tally mark in the column matching the longest uninterrupted streak.

STREAK of:	1	2	3	4	5	6	7
Results from Simulation							
Relative Frequency (nearest hundredth)							

6. Compare your prediction from # 4 with the experimental results recorded in # 5. Give reasons for any discrepancies.

7. Complete the table after running the STREAKS7 program simulating 500 trials.

STREAK of:	1	2	3	4	5	6	7	Total
Your results from running the program simulation								
Relative frequency of the outcome (nearest hundredth)								

8. Complete the table after all students have shared their results from the STREAKS7 program.

STREAK of:	1	2	3	4	5	6	7	Total
Total results from all trials from all students								
Relative frequency of the outcome (nearest hundredth)								

9. Compare your predictions to the experimental results in #7 and #8. Discuss any surprising discrepancies.

10. Record your results after analyzing the Sample Space for 7 Coin Tosses chart.

STREAK of:	1	2	3	4	5	6	7	Total
Number of Streaks								
Theoretical Probability (nearest %)								

11. Explain why the total of the theoretical probabilities is not 100%.

12. How well did the relative frequencies resulting from the many trials during the simulation predict the theoretical probabilities for each of the longest runs?

13. Examine your organized list. What is the theoretical probability of getting

a. a longest heads streak of one?

b. a longest streak of one?

Compare these probabilities with your predictions in # 3.

Teacher Notes

Activity 7

Gambler's Fallacy: Longest Streaks

Objective

- Explain the relationship between relative frequency and theoretical probability
- Confront common misconceptions about probabilities of streaks

Materials

- TI-84 Plus/TI-83 Plus

Teaching Time

- 60 minutes

Preparation

This activity continues to confront misconceptions that many people have about probability. Research suggests that the following instructional sequence helps students to address and rectify these misconceptions: first students predict results of an experiment; next they carry out the experiment or run a simulation. Students then compare the results of their experiment or simulation with their predictions, noting any discrepancies. Next, students produce a conceptual representation to represent the theoretical probability. Students may use an organized list, a tree diagram, or an area model to show all the possible outcomes and the favorable outcomes. Finally, students formulate symbolic representations or formulas to provide mathematical efficiency. If possible, students should derive, not merely be given the formulas. However, this is often not feasible with high school students.

This activity, which explores the probability of longest runs in seven coin tosses, follows this sequence. Three misconceptions that are apt to be confronted are (1) that longest runs of two occur more often than longest runs of three; (2) that longest runs of one occur frequently, and (3) that longest runs of seven out of seven tosses are extremely rare and may only occur once in 1,000 trials.

Begin the activity by distributing the Student Worksheet instead of the Exploration. Allow student only 5 - 7 minutes to respond to # 1 - 4. Students are to respond quickly based on their naive understandings; they are not to calculate exact probabilities should they have that knowledge. Engage in an in-depth classroom discussion asking students to give reasons for their predictions. It is important that students explain their thinking about their answers. You should not expect "right" answers, nor should you expect that the discussion will change students' understandings about probability. Notes are included on the answer sheet o help guide the conversation.

Next, ask students for suggestions for a simulation design for the experiment in # 4. Distribute the Exploration, which will guide students through one such simulation design. Each student will use the Probability Simulation App to gather data from 30 sets of seven tosses. It is optional whether you will have them use the program STREAKS7 (# 8 on the Student Worksheet) to gather data from 500 or more sets of seven tosses. However, the more data students have to pool for the class totals, the more they will find the results credible and be compelled to confront their misconceptions.

The program STREAKS7 should be programmed into students' graphing calculators. You may also download the program from the TI Web site: **education.ti.com**. Click on **Activities**, click on **activity books**, and then select the book *Exploring Mathematics with the Probability Simulation Application*.

The activity closes by analyzing the organized list of the 128 possible combinations in the sample space for seven tosses. It is important to end with a class discussion that guides students in confronting discrepancies between their predictions, their experimental results, and the theoretical analysis.

Answers to Student Worksheet

1. For hand **a**, there are four ways to get four kings and an ace. However, for hand **b**, there is only one way to get the five specified cards. Thus hand **a** is four times more likely to be dealt than hand **b** (1.2544 x 10^ -8 versus 3.136 x 10^ -9). Most people know that a hand with four of a kind is much less likely than one without even two of a kind; thus will choose hand **b** as the more likely hand. They are misled in thinking that hand **b** is any random selection of cards.
2. Both **a** and **b** are equally likely. Each sequence is one of the 2^7 (128) possible combinations of coin tosses, so each has a probability of occurrence of 0.7%. Since people are familiar with the fact that a coin has a 50% chance of landing on heads, many people think that tossing an alternating heads-tails sequence is much more likely than tossing all heads. Try this with your class. Ask them to write a sequence of heads and tails that mimics 10 tosses of a fair coin. Most people will try to conserve the 50-50 balance in those 10 tosses, writing a sequence that has a total of 5 or 6 heads with runs no longer than 2.
3. There are only two ways to get a longest run of one: HTHTHTH or THTHTHT. However, there are many ways to get a longest heads run of one: sequences such as HTTTTTT, THTTTTT, TTHTTTT, and so on, as well as others, such as HTHTTHT and HTTTTTH.
4. Answers may vary. Because this is the focus of the Exploration in this activity, it is important to spend time discussing students' predictions without directly addressing any misconceptions they may articulate. The element of surprise when they compare their results to their predictions is key. Have students explain their thinking for the probabilities they predicted. Remind students that their probabilities are to sum to 100%. Tally which longest streak most students think will occur most often. Listen carefully to phrases that are often confused: longest

streak of 2 versus streaks of at least 2 versus longest streaks of 2 heads versus number of streaks of 2 in a set of seven tosses, and so on. Tally how many students gave 0% for any given outcome. If some did, ask whether they think the probability is actually 0% or just very small and thus would round to 0%. Have students think about just how small is *very small*.

5. Answers will vary.

6. Answers will vary.

7. Answers will vary.

8. Answers will vary.

9. Answers will vary.

10.

STREAK of:	1	2	3	4	5	6	7	Total
Number of Streaks	2	40	46	24	10	4	2	128
Theoretical Probability (nearest %)	2%	31%	36%	19%	8%	3%	2%	101%

11. The total of the theoretical probabilities is 101%. Actual probabilities would sum to 100%; however, these are rounded probabilities that are being added together.

12. Answers will vary.

13. a. $\frac{33}{128}$ or $\approx 26\,\%$. The satisfying combinations in the order they appear in the sample space chart are: HTHTHTH, TTHTHTH, THTTHTH, HTTTHTH, TTTTHTH, THTHTTH, HTTHTTH, TTTHTTH, HTHTTTH, TTHTTTH, THTTTTH, HTTTTTH, TTTTTTH, THTHTHT, HTTHTHT, TTTHTHT, HTHTTHT, TTHTTHT, THTTTHT, HTTTTHT, TTTTTHT, HTHTHTT, TTHTHTT, THTTHTT, HTTTHTT, TTTTHTT, THTHTTT, HTTHTTT, TTTHTTT, HTHTTT, TTHTTTT, THTTTTT, HTTTTTT

b. $\frac{2}{128}$ or $\approx 2\,\%$. The satisfying combinations are HTHTHTH and THTHTHT.

Extension

The sample space for this problem consists of 3^7 or 2,187 combinations. Because the authors did not explore the problem theoretically, the theoretical probability for each longest streak was not determined.

To explore this problem experimentally, you can use the Pick Marble simulation. Use the following settings: Trial Set: 7, Types: 3, Replace: Yes. In the advanced setting, set the number of each marble type to 1. To modify the program STREAKS7, change line 6 to choose a random integer from 0-2.

An experimental result from 1,500 trials: longest streaks of 1 (8%), of 2 (53%), of 3 (27%), of 4 (8%), of 5 (3%), of 6 (1%), and of 7 (almost 0%). Notice that longest streaks of 1 and 2 are much more plentiful than they were for the coin model. In the coin model, the probability for a longest streak of one was the same as that for a longest streak of seven; this is not true for the marble experiment.

Answers to Sample Space for 7 Coin Tosses

HHHHHHH 7	HHHHHOH 5	HHHHHHO 6	HHHHHOO 5
OHHHHHH 6	OHHHHOH 4	OHHHHHO 5	OHHHHOO 4
HOHHHHH 5	HOHHHOH 3	HOHHHHO 4	HOHHHOO 3
OOHHHHH 5	OOHHHOH 3	OOHHHHO 4	OOHHHOO 3
HHOHHHH 4	HHOHHOH 2	HHOHHHO 3	HHOHHOO 2
OHOHHHH 4	OHOHHOH 2	OHOHHHO 3	OHOHHOO 2
HOOHHHH 4	HOOHHOH 2	HOOHHHO 3	HOOHHOO 2
OOOHHHH 4	OOOHHOH 3	OOOHHHO 3	OOOHHOO 3
HHHOHHH 3	HHHOHOH 3	HHHOHHO 3	HHHOHOO 3
OHHOHHH 3	OHHOHOH 2	OHHOHHO 2	OHHOHOO 2
HOHOHHH 3	HOHOHOH 1	HOHOHHO 2	HOHOHOO 2
OOHOHHH 3	OOHOHOH 2	OOHOHHO 2	OOHOHOO 2
HHOOHHH 3	HHOOHOH 2	HHOOHHO 2	HHOOHOO 2
OHOOHHH 3	OHOOHOH 2	OHOOHHO 2	OHOOHOO 2
HOOOHHH 3	HOOOHOH 3	HOOOHHO 3	HOOOHOO 3
OOOOHHH 4	OOOOHOH 4	OOOOHHO 4	OOOOHOO 4
HHHHOHH 4	HHHHOOH 4	HHHHOHO 4	HHHHOOO 4
OHHHOHH 3	OHHHOOH 3	OHHHOHO 3	OHHHOOO 3
HOHHOHH 2	HOHHOOH 2	HOHHOHO 2	HOHHOOO 3
OOHHOHH 2	OOHHOOH 2	OOHHOHO 2	OOHHOOO 3
HHOHOHH 2	HHOHOOH 2	HHOHOHO 2	HHOHOOO 3
OHOHOHH 2	OHOHOOH 2	OHOHOHO 1	OHOHOOO 3
HOOHOHH 2	HOOHOOH 2	HOOHOHO 2	HOOHOOO 3
OOOHOHH 3	OOOHOOH 3	OOOHOHO 3	OOOHOOO 3
HHHOOHH 3	HHHOOOH 3	HHHOOHO 3	HHHOOOO 4
OHHOOHH 2	OHHOOOH 3	OHHOOHO 2	OHHOOOO 4
HOHOOHH 2	HOHOOOH 3	HOHOOHO 2	HOHOOOO 4
OOHOOHH 2	OOHOOOH 3	OOHOOHO 2	OOHOOOO 4
HHOOOHH 3	HHOOOOH 4	HHOOOHO 3	HHOOOOO 5
OHOOOHH 3	OHOOOOH 4	OHOOOHO 3	OHOOOOO 5
HOOOOHH 4	HOOOOOH 5	HOOOOHO 4	HOOOOOO 6
OOOOOHH 5	OOOOOOH 6	OOOOOHO 5	OOOOOOO 7