The Geometric Die

Object of the game: Accumulate the most points by the end of the fifth round
Object of the lesson: Introduction to the geometric distribution.
Time: Length of time depends upon how many questions answered and whether computer simulation is done
Assumptions: Student will have already studied discrete probability distributions and be able to calculate an expected value.

Rules:

1. Each round begins the same way with all students standing up by their desks.
2. Teacher rolls a die. If the roll is a 1, the first round is over.
3. If the roll is anything else, the standing students' score is the number of pips showing on the die.
4. Students now decide whether to remain standing or whether to sit. If the student sits, their score remains.
5. Teacher rolls the die again. If the roll is a 1, all standing students lose all their points and the round is over. If the roll is any other number, the standing students add the showing number of pips to their previous score.
6. Play continues until a 1 is rolled, or no students are still standing. This completes a "round".
7. The game ends after five rounds have been played.

Questions:

1. How many times must the die be rolled, on average, in order to get a 1?
2. What is the expected score if my strategy is to sit after n rolls?

Notes:

1. Use a very large die – like the ones used for car rearview mirror decorations (fuzzy dice).
2. This is an example of a geometric random variable. The expected number of rolls is $1/p$ where $p = 1/6$. Thus, the expected number of rolls to get a one is 6.
3. Play this with the class until the students have decided upon a personal strategy or for 10-15 minutes (or as much time as you have to devote to the game).
4. Simulate the rolls with a calculator and/or random number table, then with a computer.
5. To calculate the expected score after n rolls, it is necessary to create the probability distribution. This can be done by physically writing out every possibility or one can appeal to combinatorics. As n increases, the number of calculations quickly becomes overwhelming. This is the value of simulation.

Question #1: How many rolls do I expect before a one is rolled?

Hand simulation using a die:

1. Each student rolls a die until a 1 appears. Count the number of rolls, including the roll that resulted in a 1.
2. Do this enough times to generate a class set of at least 200 rounds of data. Students should record their initial data in an individual chart.

3. Combine class data by recording it in the class data chart.
4. Create a histogram for the class data set.
5. Below is a typical result for n = 1000. The graph was generated by Jmp-Intro©.

Hand simulation using a calculator:

1. Use calculator to generate an integer from 1 to 6. Each time an integer is generated counts as a "roll".
2. Count the number of rolls, including the roll that resulted in a 1.
3. Record and analyze data as stated in method #1.

Hand simulation using a random number table:

1. Turn to a random number table. Look at one digit at a time. If the digit is 0, 7, 8, or 9, ignore it. Otherwise each digit counts as a "roll".
2. Count the number of rolls, including the roll that resulted in a 1.
3. Record and analyze data as stated in method #1

Computer simulation:

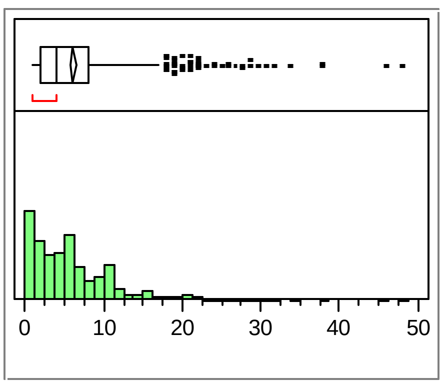
1. Open Jmp-Intro.
2. Click on **New Data Table**.
3. Double click on Column and rename it as **geom** Press enter.
4. Right click on **geom** and choose **Formula…**
5. Jmp-Intro has a built-in random number generator that follows the geometric distribution. It returns values from 0 to n, so it is necessary to add 1 to the formula. In the formula box, type **1** and click the **+** sign. Click on **Random – Random Geometric**. Replace the **p** in the red box with the fraction **1/6** (the probability of getting a "1"). Press enter. Click **OK**.
6. Click on the blue triangle to open the left side of the spreadsheet.
7. Click on the word **Untitled** and rename the worksheet to **geo data**.
8. Click on the red triangle by the worksheet title.
9. Click **New Table Property**.
10. Name it **refresh**.
11. Type **geom.<<eval formula** and click **OK**.
12. Click on the red triangle for the rows.
13. Click **Add Rows**.
14. Add as many rows as you want, up to 1000. Click **OK**. Jmp-Intro will automatically fill both the row number and the random numbers for the geometric column.
15. Click **Analyze – Distribution**.
16. Click **geom**. Click **Y, Columns**. Click **OK**.
17. You will now have a histogram and boxplot similar to the one below. Notice that there are only a few values of interest: the 5 number summary and the mean. Note also that the mean should be near 6 since that is the expected value for a geometric distribution with probability 1/6.
18. If you wish to generate a new set of data, click on the red triangle by **refresh** and click on **Run Script**. This will generate a new set of numbers that can then be analyzed the same way as before.

Individual Data Table

Round	Number of rolls
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

Class Data Table

Number of Rolls	Tally
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	



Mean		6.118
Std Dev		5.8280393
N		1000

100.0%	maximum	48.000
75.0%	quartile	8.000
50.0%	median	4.000
25.0%	quartile	2.000
0.0%	minimum	1.000

Question #2: What is the expected score if my strategy is to sit after *n* rolls?

Notes:

1. Students will first determine a personal strategy, which may not be mathematically the most advantageous, but will, instead, suit their individual personalities (risk taker, more conservative, etc.)

2. Since the expected number of rolls to get a 1 is 6, one would anticipate that the score would rise until the likelihood of rolling a 1 becomes so overpowering that the score tends back towards zero.
3. In fact, it is possible to calculate by hand the distribution of the scores for early strategies.

Analytic approach:

Strategy: Sit after n = 1 roll:

Score	**0**	**2**	**3**	**4**	**5**	**6**
Prob(score)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

A score of 0 results when a 1 is rolled. Otherwise the score is the number showing on the die. Each number is equally likely to appear, so the probabilities are all equal.

Expected score = E(score) = $\sum xp(x) = (0+2+3+4+5+6)\left(\frac{1}{6}\right) = \frac{20}{6} = \frac{10}{3}$

Thus, in the long run, a student who sits after the first roll should expect the average score for the round to be between 3 and 4 (3.33).

Strategy: Sit after n = 2 rolls:

Score	**0**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Prob(score)	$\frac{11}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

To calculate each probability, one can easily write out all the possibilities. Note that the distribution is symmetric after the score = 0. The probability for score = 0 can also be found using the geometric distribution for p = 1/6. Complete solution is shown on last page.

Expected score = E(score) = $\sum xp(x)$ =

$$0\left(\frac{11}{36}\right)+(4+12)\left(\frac{1}{36}\right)+(5+11)\left(\frac{2}{36}\right)+(6+10)\left(\frac{3}{36}\right)+(7+9)\left(\frac{4}{36}\right)+8\left(\frac{5}{36}\right)=\frac{200}{36}=\frac{50}{9}$$

So, in the long run, a student who sits after the second roll should expect the average score for the round to be between 5 and 6 (5.56).

Strategy: Sit after n = 3 rolls:

Score	**0**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
P(score)	$\frac{91}{216}$	$\frac{1}{216}$	$\frac{3}{216}$	$\frac{6}{216}$	$\frac{10}{216}$	$\frac{15}{216}$	$\frac{18}{216}$	$\frac{19}{216}$	$\frac{18}{216}$	$\frac{15}{216}$	$\frac{10}{216}$	$\frac{6}{216}$	$\frac{3}{216}$	$\frac{1}{216}$

Expected score = E(score) = $\sum xp(x) = \frac{1500}{216} = \frac{125}{18}$

Explanation for some of these probabilities is given on last page.

Although the probabilities for any strategy can be found in a similar manner, this becomes time consuming and tedious. An adequate approximation can be determined through simulation.

Hand simulation using a die:

1. Determine the particular strategy to be simulated. Call it *n*.
2. Each student rolls the die *n* times. If a 1 appears in any of the rolls, the score is 0. If the student is successful in avoiding the 1 after *n* rolls, then the total score is written down.
3. Do this enough times to generate a class set of at least 200 rounds of data. Students should record their initial data in an individual chart.
4. Combine class data by recording it in the class data chart.
5. Create a histogram for the class data set.
6. Below is a typical result for $n = 1000$. The graph was generated by Jmp-Intro©.

Hand simulation using a calculator:

1. Set up the calculator to generate *n* integers from 1 to 6. Each list of integers is a "round".
2. If any of the numbers in the roll is a 1, then the score is a zero. Otherwise the score is the sum of the integers in the round.
3. Record and analyze the data as stated in method #1.

Hand simulation using a random number table:

1. Use a random number table. Look at groups of *n* digits, being careful to use only the digits 1 through 6. Each group of digits is a "round".
2. If any of the digits in the group is a 1, then the score is zero. Otherwise the score is the sum of the digits in the group.
3. Record and analyze data as stated in method #1

Individual Data Chart

Round	**Score**
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

Class Data Chart

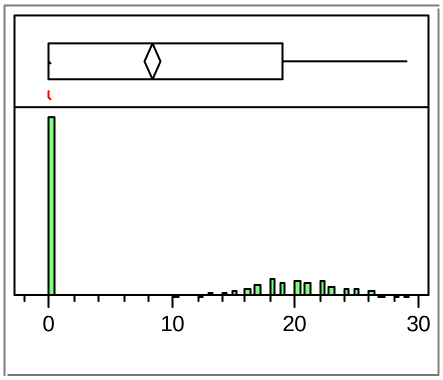
Score	**Tally**

Computer simulation:

Note: The directions are for simulating the strategy: *sit after 5 rolls*. This can be changed to any number of rolls desired by increasing or decreasing the number of columns used and adjusting the formulas appropriately.

1. Open Jmp-Intro.
2. Click on **New Data Table**.
3. Double click on **Column** and relabel the column as **int1** and press enter.
4. Double click on the next column to the right and name it **int2**. Press enter.
5. Repeat until five columns have been renamed.
6. Click on the blue triangle to open the left side of the spreadsheet.
7. Click on the name **untitled** and rename it **simulation**.
8. Jmp-Intro will create random integers for each column you have named. Right click on **int1** and click on **Formula…**
9. Click **Random – Random Integer**. Replace the **n1** in the red box with **6**. This will create random integers from 1 to 6 to represent the roll of the die. Click enter.
10. Click on the outside edge of the box to outline the entire formula in red. Press control C to copy the formula to the clipboard. You will need this formula for each of the other int columns. Click **OK**.
11. Right click each of the other columns and copy the formula into the formula box by using control V. Be certain to click **OK** after each paste.
12. Click on the red triangle for the rows.
13. Click **Add Rows**
14. Add as many rows as you want, up to 1000. Click **OK**. Jmp-Intro will automatically fill both the row number and the random numbers for each of the labeled columns.
15. Name a new column **score**.
16. Right click on the column. Click on **Formula…**
17. Click **Conditional – If**. Click **int1**, type **=** (which return a double equals sign and outlines a new box in red). Type **1** and click **enter**. Click on the outside box to outline the entire typing in red.
18. Click **Conditional – Or**. Click **int2**, type =, and click **enter**. Click on the outside box to outline everything in red.
19. Repeat until the command looks like**: *int1* = = 1 | *int2* = = 1 | *int3* = = 1 | *int4* = = 1 | *int5* = = 1**
20. Click in the **then clause** box. Type **0** and press enter.
21. Click in the **else clause** box. Click ***int1, +, int2, +, int3, +, int4, +, int5***.
22. Click **OK**. The total score is now calculated for each row of data, representing a round of play.
23. Click **Analyze – Distribution**. Click **scores**. Click **Y-Columns**.
24. Click **OK**. You will get a display similar to the one on the next page.
25. In order to run this simulation multiple times easily, click on the red triangle by the worksheet title: simulation.
26. Click **New Table Property**.
27. Name it **refresh**
28. Click in the **Value** window.

29. Type the following: **int1 << eval formula; int2 << eval formula; int3 << eval formula; int4 << eval formula; int5 << eval formula**
30. Click **OK**.
31. If you wish to generate a new set of data, click on the red triangle by **refresh** and click on **Run Script**. This will generate a new set of numbers that can be analyzed in the same manner as before. Since the exact values were not calculated (given) for this strategy, it is important to generate several sets of data to determine a better approximation for the mean score for this strategy.



100.0%	maximum	29.000
75.0%	quartile	19.000
50.0%	median	0.000
25.0%	quartile	0.000
0.0%	minimum	0.000

Mean	8.403
Std Dev	10.133572
N	1000

Complete solution for strategy: sit after 2 rolls.

p(score = 0) = $\frac{1}{6}+\frac{1}{6}\left(\frac{5}{6}\right)$ for the rolls: (1,1), (1,2), (1,3,) (1,4), (1,5), (1,6), (2,1), (3,1), (4,1), (5,1), (6,1)

p(score = 4) = $\frac{1}{36}$ for the roll: (2,2).

p(score = 5) = $\frac{2}{36}$ for the rolls: (2,3), (3,2).

p(score = 6) = $\frac{3}{36}$ for the rolls: (2,4), (3,3), (4,2)

p(score = 7) = $\frac{4}{36}$ for the rolls: (2,5), (3,4), (4,3), (5,2)

p(score = 8) = $\frac{5}{36}$ for the rolls: (2,6), (3,5), (4,4), (5,3), (6,2)

p(score = 9) = $\frac{4}{36}$ for the rolls: (3,6), (4,5), (5,4), (6,3)

p(score = 10) = $\frac{3}{36}$ for the rolls: (4,6), (5,5), (6,4)

p(score = 11) = $\frac{2}{36}$ for the rolls: (5,6), (6,5)

p(score = 12) = $\frac{1}{36}$ for the roll: (6,6)

Partial solution for the strategy: sit after 3 rolls.

p(score = 0) = $\frac{1}{6}+\frac{1}{6}\left(\frac{5}{6}\right)+\frac{1}{6}\left(\frac{5}{6}\right)^2$

p(score = 8) = $\frac{6}{216}$ found from variations of (2, 2, 4) and (2, 3, 3)

To determine variations of (2,2,4) = $\binom{3}{2}\binom{1}{1}=3$

To determine variations of (2,3,3) = $\binom{3}{1}\binom{2}{2}=3$

p(score = 11) = $\frac{18}{216}$ found from variations of (2,3,6), (2,4,5), (3,4,4), and (3,3,5)

Variations of (2,3,6) = $\binom{3}{1}\binom{2}{1}\binom{1}{1}=6$

Variations of (2,4,5) = $\binom{3}{1}\binom{2}{1}\binom{1}{1}=6$

Variations of (3,4,4) = $\binom{3}{1}\binom{2}{2}=3$

Variations of (3,3,5) = $\binom{3}{2}\binom{1}{1}=3$

All other probabilities for this table are calculated in a similar manner.