Geometric Distributions

Name ________________________

Class ________________________

Problem 1 – Introduction

In a **Geometric Distribution**, the random variable counts the number of times until a success happens. It has the following properties.

1. A trial has two options either success or failure. *P*(success) = *p*
2. Each trial is independent.
3. The variable of interest is the number of trials required to obtain the first success.

Discuss the following. Determine why each variable is or is not a geometrically distributed.

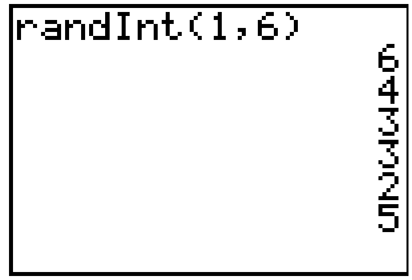
- The number of children in a family when the first boy is born.
- The expected number of boys in a family with *n* children.
- The number of coin tosses when the first heads appears.
- The number of random dart throws it takes to hit the bulls-eye.
- The number of cards drawn from a standard deck (without replacement) when the first Queen appears.
- The number of doubles rolled when rolling a pair of dice 10 times.

Problem 2 – Simulation

What is the probability that it takes four dice rolls to roll the first 5?

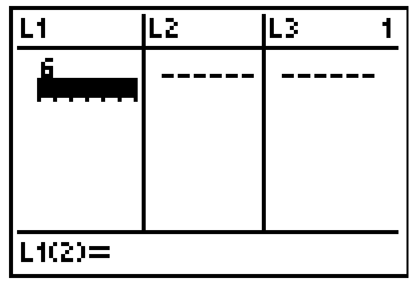
Step 1: To investigate this, enter **randInt(1, 6)** until the number 5 appears. Count the number of rolls it took and record it in list **L1**. Repeat for a total of 10 simulations.

Note: To access the **randInt** command, press [MATH], arrow to the PRB menu and choose **randInt(**.



Record your results in L1.

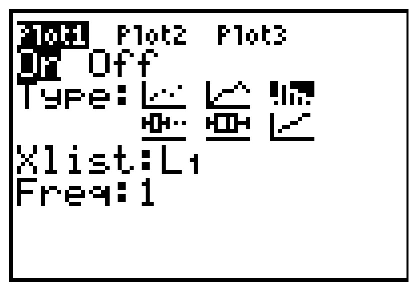
- What is your experimental probability that the first 5 appeared on the fourth roll?



Step 2: Record the numbers of the other member's of your group with the numbers in your list.

Step 3: Then create a histogram of the data. Choose [2nd] [STAT PLOT] and select **Plot1**. Match the settings as shown at the right.

Press [WINDOW] and change the settings to
Xmin = 0, Xmax = 15, Xscl = 1,
Ymin = 0, Ymax = 10, Yscl = 1.



You may use the List or the graph to determine the experimental probability of the group. Use the Trace feature and the left and right arrows to navigate the graph and determine values.

- What is your group's experimental probability that the first 5 appeared on the fourth roll?

- Using the graph, what can you determine about the number of rolls it takes for the first 5 to appear?

Problem 3 – Investigation

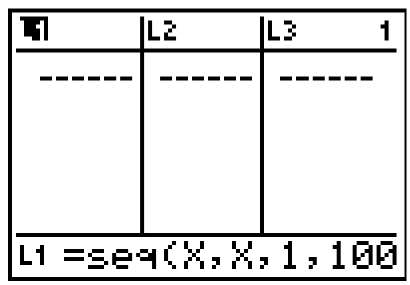
What is the theoretical probability that the first 5 appears on the fourth roll?

To investigate this, first find the theoretical probability of the first 5 appearing on the first roll, then the second, and so on. Complete the following chart. Remember that for each trial, the probability of rolling a 5 is $\frac{1}{6}$ and the probability of not rolling a 5 is $\frac{5}{6}$.

Number of Rolls	Exact Probability	Approximation
1		
2		
3		
4		
10		

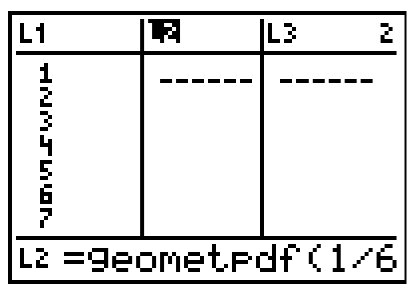
- What is the probability of the first 5 appearing on the *n*th roll?

Use lists to determine the probability of the first 5 appearing on rolls 1 through 100.



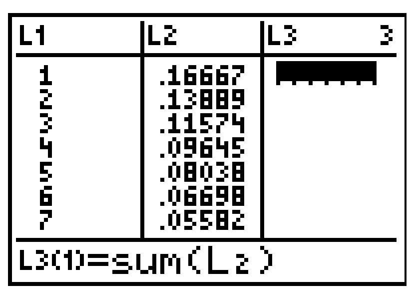
Step 4: List **L1** is the number of trials, *n*. Highlight L1, press [2nd] [LIST], arrow to the OPS menu and choose **seq(**. Then type **x,x,1,100)** and press [ENTER]. This generates a list of the integers from 1 to 100 in L1.

List **L2** is the probability that the first success will occur in *n* trials.



Step 5: Highlight L2, press [2nd] [DISTR] and choose **geometpdf(**. Then type **1/6, L1)** and press [ENTER]. This generates the probabilities.

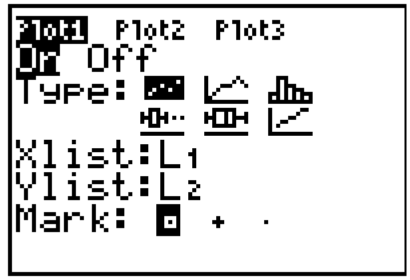
Use the first cell in L3 to display the sum of the probabilities in list L2.



Step 6: Highlight that cell, press [2nd] [LIST], arrow to the MATH menu, and choose **sum(**. Then enter L2 (by pressing [2nd] [L2]).

- How do your calculations in the table on this worksheet compare to the values in the list?

Step 7: You can also make a scatter plot of probabilities versus number of trials. Change **Plot1** to match the settings shown at the right.



- What conclusions can you draw about the probability of the first 5 appearing as the number of rolls increases?

Step 8: Find an exponential regression for the data in the scatter plot. To find the regression formula, press [STAT], arrow to the CALC menu and choose **ExpReg**. Then enter the names of the two lists, and press [ENTER].

Press [STO▸] and enter **Y1** to graph the equation.

- How does the regression equation relate to your theoretical probability? Show how you know.

- In this activity the probability of success was $p = \frac{1}{6}$. What is the general formula for any value of p?

Problem 4 – Expected Value

The expected value of a distribution is defined as the sum of the probability an outcome multiplied by the payoff. In this example, the payoff is the number of rolls.

Use the lists created in Problem 3 to calculate the expected value of this probability distribution. Calculate the products in **L3** and the sum in **L4**.

- What is the expected value? What does it represent in the context of the problem?

Extension – More or less

1. What is the probability that it will take *less than* 4 rolls to obtain a 5?

2. What is the probability that it will take *more than* 4 rolls to obtain a 5?

