Geometric Distributions

ID: 12102

Time required
45 minutes

Activity Overview

In this activity, students will simulate a geometric distribution of rolling a die until a 6 appears. They will determine experimental probabilities and draw conclusions about the number of rolls it takes for the first 6 to appear. Students will then use ***GeomPdf*** *to calculate the theoretical probabilities and use a scatter plot to draw conclusions. They will also calculate expected value.*

Topic: Random Distributions

- *Geometric Distributions*
- *Cumulative Distribution*
- *Probability*
- *Simulations*

Teacher Preparation and Notes

- *This lesson is intended to be teacher-led to help students derive the formulas for geometric distributions and the expected value. Two extension problems are included at the end of the activity. The student worksheet is used to help students take notes and draw conclusions.*
- *Students should know binomial distributions to answer the first three questions. They also need to know combinations and basic probability rules to derive the formulas.*
- ***To download the student and solution TI-Nspire documents (.tns files) and student worksheet, go to education.ti.com/exchange and enter "12102" in the quick search box.***

Associated Materials

- *GeoDist_Student.doc*
- *GeoDist.tns*
- *GeoDist_Soln.tns*

Suggested Related Activities

To download any activity listed, go to education.ti.com/exchange and enter the number in the quick search box.

- *Geometric Dartboards (TI-Nspire technology) — 8269*
- *Binomial and Geometric Distributions (TI-84 Plus and TI-Navigator) — 8395*
- *Geometric Distribution (TI-84 Plus and TI-Navigator) — 1956*

Problem 1 – Introduction

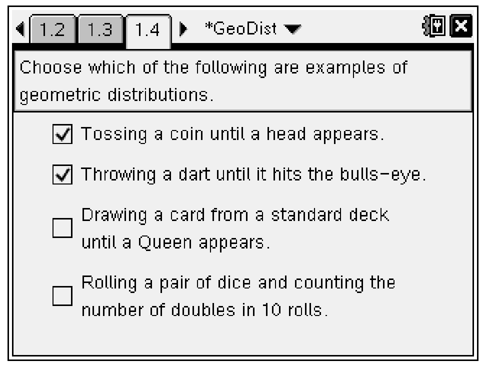
Introduce the definition for a geometric distribution. Contrast this definition with the definition for a binomial distribution.

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 with students the answers to the questions on pages 1.2 to 1.4 in the TI-Nspire document and why each situation is or is not a geometric distribution.

Drawing a card from a deck is not an independent event. Rolling dice for 10 rolls is a fixed number of trials.

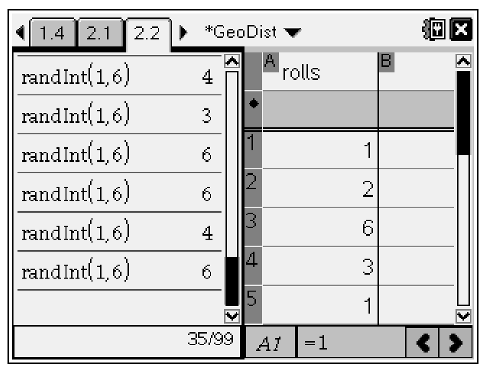


Problem 2 – Simulation

This activity centers around the geometric distribution of rolling a die until the number 6 appears. Students will use a simulation to find the number of rolls it takes to obtain the first 6.

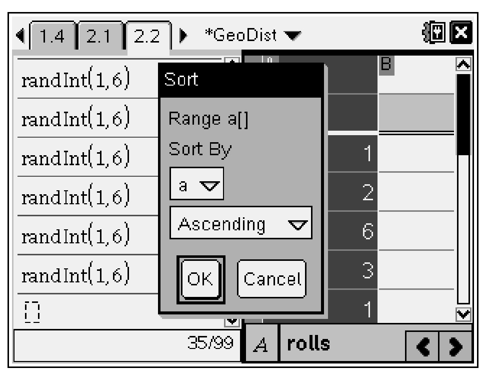
Step 1: On the calculator application, students are to enter **randInt(1,6)** and then press enter until a 6 appears.

Step 2: They need to move to the spreadsheet and record the number of rolls or number of times enter was pressed.

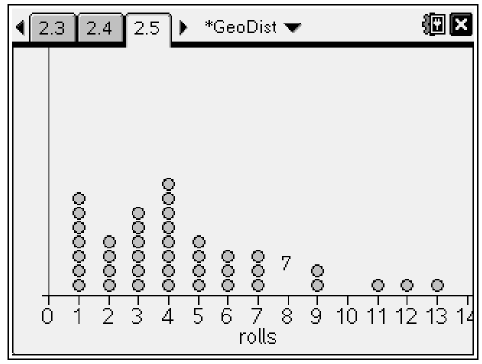


Students need to repeat steps 1 and 2 until they have completed 10 simulations. They can press **MENU > Actions > Clear History** to clear the screen after each simulation if needed.

Step 3: Students are to determine their experimental probability by sorting Column A. With Column A highlighted, press **MENU > Actions > Sort** and press enter.



Step 4: In groups of four, students are to compile all of the results by adding to list in Column A of the spreadsheet on page 2.2.

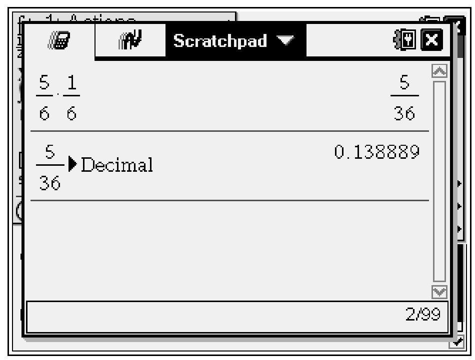


Step 5: A dot plot is shown on page 2.5. Students can use the spreadsheet or the graph to determine the experimental probability of the group.

Discuss the results of the different groups. Ask students what conclusions they can draw from the graph. They should be able to see that the first 6 appears for low numbers of rolls and the graph is skewed left. Allow this to lead into a discussion of the theoretical probability.

Problem 3 – Investigation

The goal of this part is to derive the formula for *P*(*X*=*n*), where *n* is the number of trials it takes for the first success.



Step 1: Students are to use the *Scratchpad* to calculate the probabilities of the first 6 appearing on the 1st, 2nd, 3rd, 4th, and 10th roll. (Note that in the solution document, these calculations appear on page 3.2.)

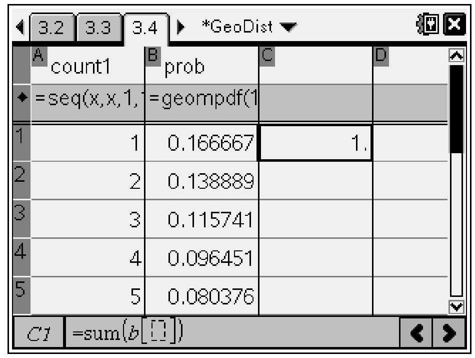
From these calculations, students should see a pattern leading them to the formula for calculating the probability of the first 6 appearing on the *n*th roll.

$P(x = 1) = \frac{1}{6}$

$P(x = 2) = \frac{5}{6} \cdot \frac{1}{6}$ (Failure, then success)

$P(x = 3) = \left(\frac{5}{6}\right)^2 \frac{1}{6}$ (2 failures, then success)

$P(x = n) = \left(\frac{5}{6}\right)^{n-1} \frac{1}{6}$ (*n* – 1 failures, then success)



Step 2: Column A is the number of trials. When students enter **=seq(x,x,1,100,1)** in the formula cell, it will generate the numbers 1 through 100.

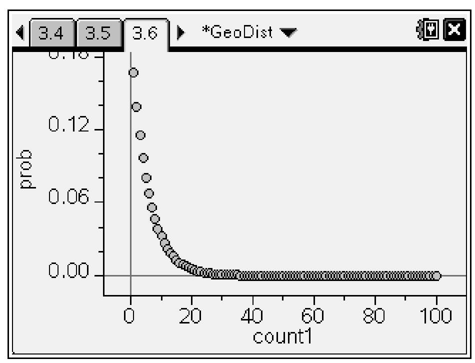
Step 3: Column B is the probability that the first success will occur in *n* trials. When students enter **=geompdf(1/6, a[])**, it will generate the probabilities. Cell C1 shows the sum of the probabilities in column B.

Discuss with students why the value in cell C1 is equal to 1 and what would happen to that number if the probabilities for rolls greater than 100 were found.

This spreadsheet will give the geometric probabilities for each number in column A when the $P(S) = \frac{1}{6}$.

Discussion Questions:

- How can you verify that this is a probability distribution?
- Theoretically, is the distribution finite or infinite?
- Practically, how long does it take before the $P(X=n) \approx 0$?

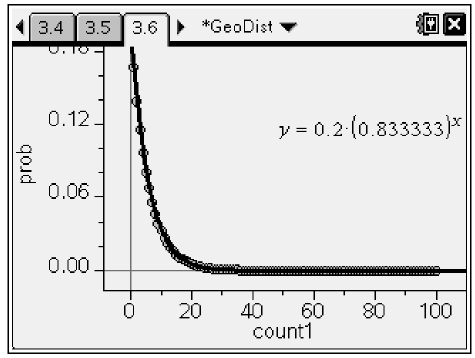


Step 4: A scatter plot of the probabilities is automatically graphed on page 2.7. Discuss with students that this is a graphical representation of the geometric distribution. This graph should reaffirm to students that the first 6 will most likely appear for a low number of rolls.

Discussion Questions:

- What is the shape of the graph?
- What type of function is represented?

Step 5: Students will now add an exponential regression line to the graph. (**MENU > Analyze > Regression > Show Exponential**)



Through some algebra steps, students should see that the regression equation is the same as the formula for the first 6 appearing on the *n*th roll.

$$\left(\frac{5}{6}\right)^{n-1}\cdot\frac{1}{6}=\left(\frac{5}{6}\right)^{-1}\cdot\frac{1}{6}\cdot\left(\frac{5}{6}\right)^{n}=\frac{6}{5}\cdot\frac{1}{6}\cdot\left(\frac{5}{6}\right)^{n}$$

$$=\frac{1}{5}\cdot\left(\frac{5}{6}\right)^{n}=0.2\left(8.3333333^{n}\right)$$

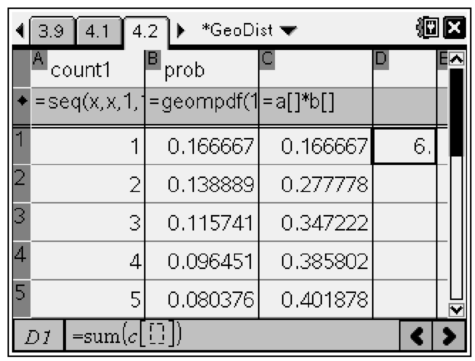
<u>Discussion Questions:</u>

- Why is this distribution called a geometric distribution? (Show the ties to a geometric sequence. This can also be used to **prove** that the sum of a geometric distribution is 1.)
- How can this formula be made into a general formula for geometric distributions?

$$P(X = n) = (1 - p)^{n-1}\, p$$

Problem 4 – Expected Value

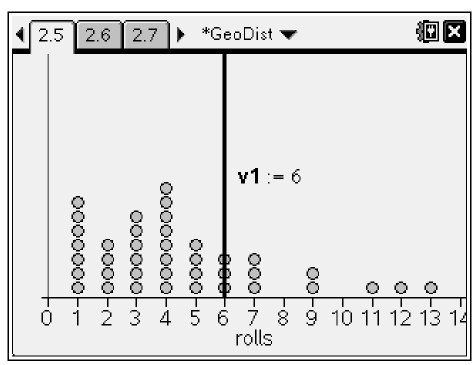
Students will use the spreadsheet on page 4.2 to calculate the expected value. Columns A and B from the previous spreadsheet have already been entered. Discuss with students the definition of expected value of a random distribution.



Step 1: Students are to enter **= a[ ] * b[ ]** in the formula bar of Column C.

Step 2: In cell D1, they need to enter **=sum(c[ ])**.

Students should understand that the expected value of 6 represents the average number of rolls until the first 6 appears. They can plot this number on the graph back on page 2.5 to see that it is approximately the mean of the distribution (**MENU > Analyze > Plot Value**).



<u>Discussion Questions:</u>

- Is this answer reasonable?
- How can the general formula for expected value be found? (This is can be done using a geometric series).

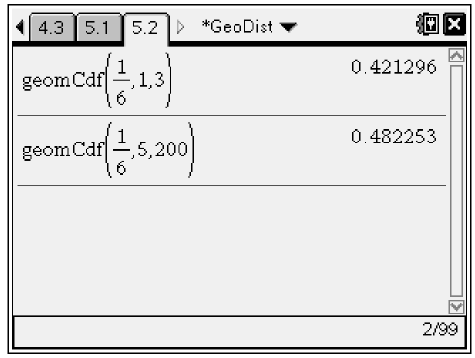
In general, the expected value of a geometric distribution is $\frac{1}{p}$.

Extension – More or less

Students are asked two questions that involve adding geometric probabilities.

What is the probability that it will take less than 4 rolls to obtain a six?

Solution: $P(X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$



This can also be found by using the **geomCdf** command. On page 5.2, students should press **MENU > Probability > Distributions > Geometric Cdf**. A drop down menu will appear asking for *p* (probability of success), lower bound, and upper bound.

What is the probability that it will take more than 4 rolls to obtain a six?

Solution: $P(X > 4)$ = geomCdf(1/6,5,200) (Note: A large number was chosen as the upper bound to ensure all the probabilities greater than 0 were added.)

