A geometric distribution describes the probability of waiting x trials for the first success when each trial has p chance of success.

A general geometric distribution

x	$P(X = x)$
1	p
2	$(1 - p)p$
3	$(1 - p)^2 p$
4	$(1 - p)^3 p$
5	$(1 - p)^4 p$
6	$(1 - p)^5 p$
$\vdots$	$\vdots$
n	$(1 - p)^{n-1} p$
$\vdots$	$\vdots$

We may use the abbreviation $\mathsf{Geo}(p)$. So, for example, if $Y \sim \mathsf{Geo}(0.4)$, we can show the distribution in a table.

Distribution of $Y \sim \mathsf{Geo}(0.4)$

y	$P(Y = y)$
1	0.4
2	0.24
3	0.144
4	0.0864
5	0.05184
6	0.031104
$\vdots$	$\vdots$
n	$(0.6)^{n-1}(0.4)$
$\vdots$	$\vdots$

For geometric distributions, the probabilities are always decreasing. We can't show every possible outcome because there are infinite possibile outcomes.

Some people who loved math figured out that even though there are infinite possibilities, we can still characterize this distribution with a mean and standard deviation.

For $\mathsf{Geo}(p)$:

$$\mu = \frac{1}{p}$$

$$\sigma = \frac{\sqrt{1 - p}}{p} = \sqrt{\frac{1 - p}{p^2}}$$

1: Let $X \sim \text{Geo}(p = 0.7)$.

    a: Complete the first few rows of the probability distribution table.

x	$P(X = x)$
1	
2	
3	
4	
5	
$\vdots$	$\vdots$
n	
$\vdots$	$\vdots$

    b: Calculate the probability $P(X \leq 3)$.

    c: Calculate the probability $P(X > 3)$.

    d: Estimate the mean by evaluating the following.

$$\sum_{x=1}^{5} x \cdot P(X = x)$$

    e: Calculate the mean using the formula for geometric distributions.

2: Please roll a 6-sided die until you roll a 6. Please count how many trials it takes. For example, if you rolled [3,2,4,4,6], it took 5 trials to hit 6 and you just record the 5.

3: Repeat the above process 9 more times. Record the number of trials it takes each time. Your instructor will ask for your numbers to make a histogram of all the class's results. Overall, you'll expect a list of 10 numbers, kind of like

$$5 \quad 12 \quad 7 \quad 2 \quad 8 \quad 1 \quad 4 \quad 10 \quad 1 \quad 8$$

4: Find the sample mean and sample standard deviation from your 10 numbers. Also, tell your instructor your sample mean and sample standard deviation; we will make another histogram of these as a preview of *sampling distributions* and the *central limit theorem.*

5: Let $X \sim \text{Geo}\left(\frac{1}{6}\right)$. Calculate μ and σ.

6: Let $X \sim \mathcal{N}(200,\ 6)$. Calculate $P(X \geq 210)$.

7: Let $X \sim \mathsf{Geo}(p = 0.2)$. Calculate $P(X \leq 2)$.

8: Determine x_0 such that $X \sim \mathcal{N}(50,\ 0.2)$ and $P(X > x_0) = 0.75$.