Discrete Random Variables: Geometric (optional)*

Susan Dean
Barbara Illowsky, Ph.D.

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 3.0†

Abstract

This module describes the geometric experiment and the geometric probability distribution. This module is included in the Collaborative Statistics textbook/collection as an optional lesson.

The characteristics of a geometric experiment are:

1. There are one or more Bernoulli trials with all failures except the last one, which is a success. In other words, you keep repeating what you are doing until the first success. Then you stop. For example, you throw a dart at a bull's eye until you hit the bull's eye. The first time you hit the bull's eye is a "success" so you stop throwing the dart. It might take you 6 tries until you hit the bull's eye. You can think of the trials as failure, failure, failure, failure, failure, success. STOP.
2. In theory, the number of trials could go on forever. There must be at least one trial.
3. The probability,p, of a success and the probability, q, of a failure is the same for each trial. $p + q = 1$ and $q = 1 - p$. For example, the probability of rolling a 3 when you throw one fair die is $\frac{1}{6}$. This is true no matter how many times you roll the die. Suppose you want to know the probability of getting the first 3 on the fifth roll. On rolls 1, 2, 3, and 4, you do not get a face with a 3. The probability for each of rolls 1, 2, 3, and 4 is $q = \frac{5}{6}$, the probability of a failure. The probability of getting a 3 on the fifth roll is $\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} = 0.0804$

$X =$ the number of independent trials until the first success. The mean and variance are in the summary in this chapter.

Example 1

You play a game of chance that you can either win or lose (there are no other possibilities) **until** you lose. Your probability of losing is p = 0.57. What is the probability that it takes 5 games until you lose? Let $X =$ the number of games you play until you lose (includes the losing game). Then X takes on the values 1, 2, 3, ... (could go on indefinitely). The probability question is $P(x = 5)$.

Example 2

A safety engineer feels that 35% of all industrial accidents in her plant are caused by failure of employees to follow instructions. She decides to look at the accident reports (selected randomly and replaced in the pile after reading) **until** she finds one that shows an accident caused by failure

*Version 1.16: May 27, 2012 4:02 am -0500
†http://creativecommons.org/licenses/by/3.0/

of employees to follow instructions. On the average, how many reports would the safety engineer **expect** to look at until she finds a report showing an accident caused by employee failure to follow instructions? What is the probability that the safety engineer will have to examine at least 3 reports until she finds a report showing an accident caused by employee failure to follow instructions?

Let X = the number of accidents the safety engineer must examine **until** she finds a report showing an accident caused by employee failure to follow instructions. X takes on the values 1, 2, 3, ... The first question asks you to find the **expected value** or the mean. The second question asks you to find $P(x \geq 3)$. ("At least" translates as a "greater than or equal to" symbol).

Example 3

Suppose that you are looking for a student at your college who lives within five miles of you. You know that 55% of the 25,000 students do live within five miles of you. You randomly contact students from the college **until** one says he/she lives within five miles of you. What is the probability that you need to contact four people?

This is a geometric problem because you may have a number of failures before you have the one success you desire. Also, the probability of a success stays the same each time you ask a student if he/she lives within five miles of you. There is no definite number of trials (number of times you ask a student).

Problem 1

Let X = the number of _____ you must ask _____ one says yes.

Solution

Let X = the number of **students** you must ask **until** one says yes.

Problem 2 ***(Solution on p. 4.)***

What values does X take on?

Problem 3 ***(Solution on p. 4.)***

What are p and q?

Problem 4 ***(Solution on p. 4.)***

The probability question is P(_____).

1 Notation for the Geometric: G = Geometric Probability Distribution Function

$X \sim \mathrm{G}(\mathrm{p})$

Read this as "X is a random variable with a geometric distribution." The parameter is p. p = the probability of a success for each trial.

Example 4

Assume that the probability of a defective computer component is 0.02. Components are randomly selected. Find the probability that the first defect is caused by the 7th component tested. How many components do you expect to test until one is found to be defective?

Let X = the number of computer components tested until the first defect is found.

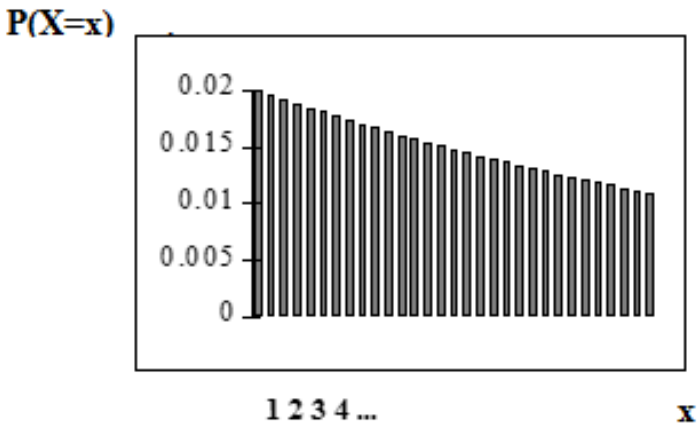
X takes on the values 1, 2, 3, ... where $p = 0.02$. $X \sim \mathrm{G}(0.02)$

Find $P(x = 7)$. $P(x = 7) = 0.0177$. (calculator or computer)

TI-83+ and TI-84: **For a general discussion, see this example (binomial)**. The syntax is similar. The geometric parameter list is (p, number) If "number" is left out, the result is the geometric probability table. For this problem: **After you are in 2nd DISTR, arrow down to D:geometpdf. Press ENTER. Enter .02,7). The result is $P(x = 7) = 0.0177$.**

The probability that the 7th component is the first defect is 0.0177.

The graph of $X \sim \mathrm{G}(0.02)$ is:



The y-axis contains the probability of x, where $X =$ the number of computer components tested.

The number of components that you would expect to test until you find the first defective one is the mean, $\mu = 50$.

The formula for the mean is $\mu = \frac{1}{p} = \frac{1}{0.02} = 50$

The formula for the variance is $\sigma^2 = \frac{1}{p} \cdot \left(\frac{1}{p} - 1\right) = \frac{1}{0.02} \cdot \left(\frac{1}{0.02} - 1\right) = 2450$

The standard deviation is $\sigma = \sqrt{\frac{1}{p} \cdot \left(\frac{1}{p} - 1\right)} = \sqrt{\frac{1}{0.02} \cdot \left(\frac{1}{0.02} - 1\right)} = 49.5$

Solutions to Exercises in this Module

Solution to Example 3, Problem 2 (p. 2)
1, 2, 3, ..., (total number of students)
Solution to Example 3, Problem 3 (p. 2)

- $p = 0.55$
- $q = 0.45$

Solution to Example 3, Problem 4 (p. 2)
$P(x = 4)$

Glossary

Definition 1: Geometric Distribution
A discrete random variable (RV) which arises from the Bernoulli trials. The trials are repeated until the first success. The geometric variable X is defined as the number of trials until the first success. Notation: $X \sim G(p)$. The mean is $\mu = \frac{1}{p}$ and the standard deviation is $\sigma = \sqrt{\frac{1}{p} \cdot \left(\frac{1}{p} - 1\right)}$ The probability of exactly x failures before the first success is given by the formula: $P(X = x) = p(1-p)^{x-1}$.