

# Learning the Geometric Distribution: A Practical Guide with Excel Examples

Authored by  
**Mohammed loot**

November 3, 2025

## RECOMMENDED CITATION

Mohammed loot (2025). *Learning the Geometric Distribution: A Practical Guide with Excel Examples*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=9434>

The [Geometric Distribution](#) is a fundamental concept in sequential [probability](#) theory, specifically designed to model the number of independent trials required until the very first success is achieved. This unique focus makes the distribution indispensable for understanding waiting times in processes where outcomes are strictly binary and the trials are independent.

In practical statistical modeling, the geometric distribution describes the probability of experiencing a certain amount of **failures** before finally experiencing the first **success** in a series of repeated statistical experiments. Mastering the calculation of this distribution, particularly within a powerful software environment like [Excel](#), allows professionals to perform quick and accurate statistical predictions across diverse fields, including actuarial science, quality control, and engineering reliability analysis.

A [Bernoulli trial](#) is formally defined as a random experiment that yields only two possible, mutually exclusive outcomes--conventionally labeled "success" or "failure." Crucially, for a sequence of trials to be modeled geometrically, the probability of success, denoted by  $p$ , must remain rigorously constant for every single repetition of the experiment.

A classic and easily visualized example of a [Bernoulli trial](#) is the simple flip of a fair coin. The result can only be heads (defined here as a "success") or tails (a "failure"), and the probability of success (0.5) must remain the same for each flip, provided the coin is fair and each flip is genuinely independent of the last.

When a [random variable](#)  $X$  is said to follow a geometric distribution, the associated probability mass function (PMF) calculates the likelihood of needing exactly  $k$  failures before achieving the first success. It is vital to note that this is distinct from--though related to--calculating the probability that the first success occurs specifically on the  $k+1$  trial.

## Deconstructing the Geometric Distribution Formula

To calculate the probability of observing exactly  $k$  consecutive failures prior to the first successful outcome, we employ the probability mass function (PMF). This calculation relies on the principle of multiplying the probabilities of independent events occurring in a specific sequence.

Specifically, we must account for  $k$  failures, each with a probability of  $(1-p)$ , and then multiply by the probability  $p$  for the final, successful trial. The resulting formula is concise and powerful:

$$P(X=k) = (1-p)^k p$$

The variables within this fundamental formula are defined precisely as follows:

**k:** Represents the exact number of consecutive **failures** observed immediately before the first

success occurs (where  $k$  is a non-negative integer:  $k = 0, 1, 2, \dots$ ).

**p:** Represents the constant [probability](#) of **success** on any single trial in the sequence ( $p$  must satisfy  $0 < p \leq 1$ ).

**(1-p):** Represents the constant probability of failure on any single trial. This term is often referred to as  $q$ .

The term  $(1-p)^k$  mathematically captures the probability of the necessary sequence of  $k$  preceding failures. By multiplying this term by  $p$ , we complete the sequence, finding the exact probability of this specific waiting time. The examples presented below demonstrate how to efficiently calculate these specific probabilities using the dedicated, built-in statistical functions available in [Excel](#).

## Leveraging the GEOM.DIST Function in Excel

While manually calculating the geometric probability using the PMF formula  $P(X=k) = (1-p)^k p$  is possible, [Excel](#) provides a dedicated and robust function,

GEOM.DIST

, to handle these calculations with precision and ease. This function is essential for analysts who need rapid statistical feedback without the risk of manual calculation errors.

The primary syntax for the

GEOM.DIST

function is structured to accept the key parameters of the distribution:

GEOM.DIST(number\_f, probability\_s, cumulative)

. A thorough understanding of each argument is critical for generating accurate results and distinguishing between the PMF and Cumulative Distribution Function (CDF).

The arguments are defined to align directly with the core parameters of the geometric model:

**number\_f (k):** This input specifies the exact number of consecutive failures before the first success occurs. It must always be an integer value greater than or equal to 0.

**probability\_s (p):** This is the constant [probability](#) of success assumed for any single trial within the sequence.

**cumulative:** This is a logical argument that dictates the type of probability returned. If set to **FALSE**, Excel calculates the exact geometric probability ( $P(X=k)$ , the PMF). If set to **TRUE**, it calculates the cumulative probability ( $P(X \leq k)$ ), which represents the chance that the first success

occurs after  $k$  or fewer failures.

Since the objective of analyzing the geometric distribution is often to determine the probability of a precise sequence length (i.e., exactly  $k$  failures), the

cumulative

argument for all the following examples will be set strictly to **FALSE**. This ensures we are always calculating the probability mass function for a specific outcome.

## Case Study 1: Analyzing Success Rates in a Coin Flip Simulation

Imagine a researcher is simulating a sequence of flips using a perfectly fair coin. We are interested in determining the exact [probability](#) that the sequence will require exactly three "failures" (tails) until the coin finally lands on heads (designated as the first success). This scenario perfectly fits the geometric distribution model.

In this specific case, the parameters are easily defined: we have  $k = 3$  (representing three sequential failures) and  $p = 0.5$  (the constant probability of success on any single flip). We must use the

GEOM.DIST

function in [Excel](#) to calculate the probability of the sequence Failure, Failure, Failure, Success (T, T, T, H).

GEOM.DIST(3, 0.5, FALSE)

This calculation yields the probability of this exact, sequential outcome. It is equivalent to calculating  $(1-0.5)^3 * 0.5$ . The visual output in Excel confirms the structure and result of this simple simulation.

|    | A                                           | B                 | C | D | E |
|----|---------------------------------------------|-------------------|---|---|---|
| 1  | Prob. of success on given trial (p)         | 0.5               |   |   |   |
| 2  | Number of failures before first success (k) | 3                 |   |   |   |
| 3  |                                             |                   |   |   |   |
| 4  | Geometric Distribution Formula              | $=(1-B1)^{B2}*B1$ |   |   |   |
| 5  | Probability                                 | 0.0625            |   |   |   |
| 6  |                                             |                   |   |   |   |
| 7  |                                             |                   |   |   |   |
| 8  |                                             |                   |   |   |   |
| 9  |                                             |                   |   |   |   |
| 10 |                                             |                   |   |   |   |
| 11 |                                             |                   |   |   |   |
| 12 |                                             |                   |   |   |   |
| 13 |                                             |                   |   |   |   |
| 14 |                                             |                   |   |   |   |
| 15 |                                             |                   |   |   |   |
| 16 |                                             |                   |   |   |   |
| 17 |                                             |                   |   |   |   |
| 18 |                                             |                   |   |   |   |
| 19 |                                             |                   |   |   |   |
| 20 |                                             |                   |   |   |   |
| 21 |                                             |                   |   |   |   |
| 22 |                                             |                   |   |   |   |

Based on the geometric distribution calculation performed by Excel, the probability that we will experience exactly three failures (tails) until the coin finally lands on heads is exactly **.0625**. This translates to a 6.25% chance of this specific waiting time occurring.

## Case Study 2: Calculating Free Throw Streaks in Sports

Consider a professional basketball player who maintains a consistent free throw percentage of 60%. We define a successful trial as making the shot, meaning the probability of success is fixed at  $p = 0.6$ . We are interested in finding the probability that this player will miss exactly four consecutive free throws (failures) before finally making the fifth shot (the first success).

In this application of the [Geometric Distribution](#), our parameters are defined as  $k = 4$  (four consecutive failures) and  $p = 0.6$  (probability of success). This means the probability of failure (missing the shot) is constant at  $1 - 0.6 = 0.4$ . Since each free throw is considered an independent [Bernoulli trial](#), the geometric model applies perfectly.

We apply the geometric distribution function in [Excel](#) using the required parameters, specifying **FALSE** for the cumulative argument to ensure we calculate the PMF for exactly four misses:

GEOM.DIST(4, 0.6, FALSE)

This formula models the specific, low-probability sequence: Miss, Miss, Miss, Miss, Make. The calculation determines the likelihood of this extended waiting time until success.

|    | A                                           | B                 | C | D | E | F |
|----|---------------------------------------------|-------------------|---|---|---|---|
| 1  | Prob. of success on given trial (p)         | 0.6               |   |   |   |   |
| 2  | Number of failures before first success (k) | 4                 |   |   |   |   |
| 3  |                                             |                   |   |   |   |   |
| 4  | Geometric Distribution Formula              | $=(1-B1)^{B2}*B1$ |   |   |   |   |
| 5  | Probability                                 | 0.01536           |   |   |   |   |
| 6  |                                             |                   |   |   |   |   |
| 7  |                                             |                   |   |   |   |   |
| 8  |                                             |                   |   |   |   |   |
| 9  |                                             |                   |   |   |   |   |
| 10 |                                             |                   |   |   |   |   |
| 11 |                                             |                   |   |   |   |   |
| 12 |                                             |                   |   |   |   |   |
| 13 |                                             |                   |   |   |   |   |
| 14 |                                             |                   |   |   |   |   |
| 15 |                                             |                   |   |   |   |   |
| 16 |                                             |                   |   |   |   |   |
| 17 |                                             |                   |   |   |   |   |
| 18 |                                             |                   |   |   |   |   |
| 19 |                                             |                   |   |   |   |   |
| 20 |                                             |                   |   |   |   |   |
| 21 |                                             |                   |   |   |   |   |
| 22 |                                             |                   |   |   |   |   |

The resulting probability that the player will miss four free throws until he finally makes one is calculated as **.01536**. This demonstrates that while possible, there is a relatively low 1.536% chance of this exact sequence of failures occurring before the first successful attempt.

### Case Study 3: Market Research and Sequential Sampling

Consider a market researcher conducting interviews where they ask individuals if they support a new piece of legislation. Based on preliminary data, the constant [probability](#) that a randomly selected person supports the law (success) is estimated to be  $p = 0.2$ . The researcher needs to calculate the probability that they must talk to exactly three people who do not support the law (failures) before the fourth person they interview is the first supporter.

In this sequential sampling scenario, the success is supporting the law, and the failure is not supporting it. We identify our necessary parameters: **k = 3** (three non-supporters, or failures) and **p = 0.2** (probability of success). The probability of failure, (1-p), is 0.8.

We calculate the geometric distribution probability using the following function in [Excel](#):

GEOM.DIST(3, 0.2, FALSE)

This calculation models the specific sequence of outcomes: Fail, Fail, Fail, Success. The result is derived from  $0.8^3 * 0.2$ .

|    | A                                           | B                 | C | D | E |
|----|---------------------------------------------|-------------------|---|---|---|
| 1  | Prob. of success on given trial (p)         | 0.2               |   |   |   |
| 2  | Number of failures before first success (k) | 3                 |   |   |   |
| 3  |                                             |                   |   |   |   |
| 4  | Geometric Distribution Formula              | $=(1-B1)^{B2}*B1$ |   |   |   |
| 5  | Probability                                 | 0.1024            |   |   |   |
| 6  |                                             |                   |   |   |   |
| 7  |                                             |                   |   |   |   |
| 8  |                                             |                   |   |   |   |
| 9  |                                             |                   |   |   |   |
| 10 |                                             |                   |   |   |   |
| 11 |                                             |                   |   |   |   |
| 12 |                                             |                   |   |   |   |
| 13 |                                             |                   |   |   |   |
| 14 |                                             |                   |   |   |   |
| 15 |                                             |                   |   |   |   |
| 16 |                                             |                   |   |   |   |
| 17 |                                             |                   |   |   |   |
| 18 |                                             |                   |   |   |   |
| 19 |                                             |                   |   |   |   |
| 20 |                                             |                   |   |   |   |
| 21 |                                             |                   |   |   |   |
| 22 |                                             |                   |   |   |   |
| 23 |                                             |                   |   |   |   |

The calculated probability that the fourth person the researcher talks to is the first person to support the law (following three failures) is **.1024**, equivalent to 10.24%. This provides the researcher with a quantifiable expectation for their sampling efforts.

## Summary of Key Geometric Distribution Prerequisites

The [Geometric Distribution](#) is an indispensable tool for modeling waiting times until the first

success in a sequence. Its primary strength lies in its applicability to repeated, independent [Bernoulli trials](#). By utilizing the

GEOM.DIST

function in Excel and correctly setting the

cumulative

argument to **FALSE**, users can efficiently determine the exact probability of achieving  $k$  failures before that critical first success.

It is essential to remember that the validity of the Geometric Distribution hinges on two critical assumptions that must be met in the underlying experiment: first, that all trials are statistically **independent** of one another; and second, that the **probability of success ( $p$ )** is constant and unchanging across every trial.

When these prerequisites are satisfied, the geometric distribution offers a reliable mathematical framework for analyzing success expectancy and quantifying the likelihood of specific waiting periods.

## Additional Resources for Discrete Probability

To further expand your understanding of discrete [probability](#) distributions and their application within statistical software environments, consider exploring these related topics:

The Binomial Distribution (used for modeling the total number of successes in a fixed number of trials).

The Negative Binomial Distribution (used for modeling the number of trials needed for the  $r$ -th success, a generalization of the geometric distribution).

Using statistical functions in [Excel](#) for hypothesis testing and confidence interval estimation.