

# What is a Probability Mass Function (PMF) in Statistics?

Authored by  
**Mohammed loot**

November 6, 2025

## RECOMMENDED CITATION

Mohammed loot (2025). *What is a Probability Mass Function (PMF) in Statistics?*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=11425>

The **Probability Mass Function**, universally abbreviated as the **PMF**, stands as a fundamental pillar in descriptive and inferential **statistics**. It serves the essential purpose of defining the probability distribution for a **discrete random variable**. Simply put, the PMF assigns a precise probability to every distinct, quantifiable outcome that a specific random experiment can produce. It maps the set of possible outcomes directly to their likelihoods.

It is vital to distinguish the PMF from distributions used for continuous data. Unlike continuous variables--which require a Probability Density Function (PDF) because the probability of any single exact value is zero--the PMF is exclusively designed for situations where outcomes are finite or countably infinite. This mathematical distinction makes the PMF the definitive tool for analyzing any phenomenon involving counting, such as tallying manufacturing defects, analyzing the results of a series of coin flips, or predicting the frequency of events over a specific time period.

## Understanding the PMF Through a Classic Example: Rolling a Die

To truly grasp the mechanics of the Probability Mass Function, let us consider the classic and straightforward example of rolling a standard, unbiased six-sided die one time. We define the random variable  $X$  as the specific outcome shown on the die face. Since the die is fair, each potential outcome (1, 2, 3, 4, 5, or 6) possesses an identical probability of occurring, which is  $1/6$ .

The PMF formalism allows us to describe the probability distribution of  $X$  by listing the precise probability associated with each value in the sample space. This simple model perfectly illustrates how the PMF works by providing a discrete probability value for every countable result:

$$P(X=1): 1/6$$

$$P(X=2): 1/6$$

$$P(X=3): 1/6$$

$$P(X=4): 1/6$$

$$P(X=5): 1/6$$

$$P(X=6): 1/6$$

The visual representation below provides a graphical snapshot of this distribution, demonstrating how the probability  $P(X=x)$  is mapped directly to each distinct numerical outcome  $x$  within the die's sample space. Notice that the bars, representing the mass of probability, are separated, emphasizing the discrete nature of the variable.

$$p_X(x) = \begin{cases} 1/6 & \text{if } x = 1 \\ 1/6 & \text{if } x = 2 \\ 1/6 & \text{if } x = 3 \\ 1/6 & \text{if } x = 4 \\ 1/6 & \text{if } x = 5 \\ 1/6 & \text{if } x = 6 \\ 0 & \text{otherwise} \end{cases}$$

This diagram clearly links the possible numerical outcomes on the right to their associated probability values on the left. The subsequent figure further clarifies this relationship, underscoring the definition of the [Probability Mass Function](#) as a mapping mechanism.

Probabilities	Outcomes
1/6	if $x = 1$
1/6	if $x = 2$
1/6	if $x = 3$
1/6	if $x = 4$
1/6	if $x = 5$
1/6	if $x = 6$
0	otherwise

## The Crucial Concepts of Support and Normalization

When constructing or analyzing any [PMF](#), two concepts are non-negotiable: the definition of its support and the requirement for normalization. The [support](#) of a probability mass function is formally defined as the comprehensive set of all values for which the probability is strictly positive. These are the only possible results the [discrete random variable](#) can actually adopt.

Returning to our example of the die roll, the support \$\$\$ is mathematically defined as  $S = \{1, 2, 3, 4, 5, 6\}$ . A critical implication of defining the support is that for any value  $x$  existing outside of this defined set, the Probability Mass Function must yield a probability of zero. For instance, the probability of the die landing on "0" or "7" must be  $P(X=0) = 0$  and  $P(X=7) = 0$ , respectively, since these outcomes are not included within the defined [support](#) of the experiment.

Furthermore, all valid PMFs must satisfy the fundamental axiom of probability known as

**normalization:** the sum of all probabilities assigned to every single outcome within the entire support must equal exactly 1.0. This ensures that the function accounts for 100% of all possible outcomes of the random experiment. For our six-sided die example, this condition is easily verified:

Sum of probabilities =  $1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1.0$ .

## Applying the PMF: Analyzing Named Discrete Distributions

While theoretical examples like coin flips or die rolls are excellent for foundational understanding, the practical application of PMFs most frequently occurs within the context of specific, named probability distributions. These standardized distributions provide powerful mathematical frameworks for modeling various complex, real-world phenomena that inherently involve discrete, countable outcomes. The two most ubiquitous and widely applied discrete distributions that utilize the PMF are the [Binomial Distribution](#) and the [Poisson Distribution](#).

### Case Study 1: The Binomial Distribution PMF

The [Binomial Distribution](#) is designed to model the number of observed successes ( $k$ ) resulting from a fixed number of independent trials ( $n$ ), under the condition that the probability of success ( $p$ ) remains constant across all trials. This structure is ideal for situations that can be categorized as a sequence of Bernoulli trials. If a [discrete random variable](#)  $X$  adheres to the Binomial distribution, the exact probability of achieving  $k$  successes is precisely calculated using its dedicated [PMF](#) formula:

$$P(X=k) = nCk * p^k * (1-p)^{n-k}$$

The variables within this powerful formula are defined as follows:

**n:** The total number of independent trials conducted.

**k:** The specific number of successes we are interested in calculating the probability for.

**p:** The constant probability of success occurring on any single trial.

**nCk:** The combinatorial term representing the number of unique ways to select  $k$  successes out of  $n$  trials.

As a concrete illustration, imagine we are flipping a fair coin ( $p=0.5$ ) exactly three times ( $n=3$ ). We can substitute these parameters into the Binomial PMF to determine the probabilities for observing 0, 1, 2, or 3 heads over the course of these three flips:

$$P(X=0) = 3C0 * .5^0 * (1-.5)^{3-0} = 1 * 1 * (.5)^3 = \mathbf{0.125}$$

$$P(X=1) = 3C1 * .5^1 * (1-.5)^{3-1} = 3 * .5 * (.5)^2 = \mathbf{0.375}$$

$$P(X=2) = 3C2 * .5^2 * (1-.5)^{3-2} = 3 * .25 * (.5)^1 = \mathbf{0.375}$$

$$P(X=3) = 3C3 * .53 * (1-.5)3-3 = 1 * .125 * (.5)0 = \mathbf{0.125}$$

## Case Study 2: The Poisson Distribution PMF

The [Poisson Distribution](#) is invaluable in fields such as quality control, queuing theory, and epidemiology. It specifically models the probability of a certain number of independent events ( $k$ ) occurring within a fixed, well-defined interval of time or space, provided that these events occur at a known, constant average rate ( $\lambda$ ). If the [random variable](#)  $X$  follows a Poisson process, the probability that  $X$  equals  $k$  occurrences is derived from its PMF:

$$P(X=k) = \lambda^k * e^{-\lambda} / k!$$

The terms utilized in the Poisson PMF are defined as:

$\lambda$  (**Lambda**): The mean or average number of successes (events) expected to occur during the specified interval.

$k$ : The specific, non-negative integer number of successes (events) being observed.

$e$ : The mathematical constant known as [Euler's number](#), approximately equal to 2.71828.

Consider a practical scenario where a hospital registers an average rate ( $\lambda$ ) of 2 births per hour. Using the Poisson PMF, we can calculate the probability of observing 0, 1, 2, or 3 births within any given one-hour interval:

$$P(X=0) = 2^0 * e^{-2} / 0! = \mathbf{0.1353}$$

$$P(X=1) = 2^1 * e^{-2} / 1! = \mathbf{0.2707}$$

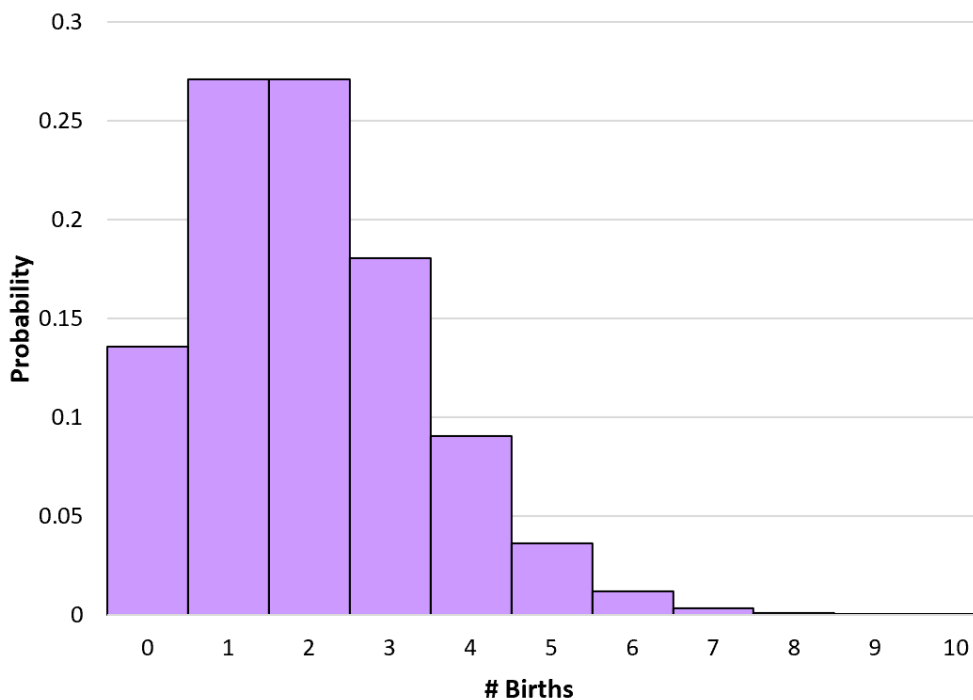
$$P(X=2) = 2^2 * e^{-2} / 2! = \mathbf{0.2707}$$

$$P(X=3) = 2^3 * e^{-2} / 3! = \mathbf{0.1805}$$

## Visualizing the Probability Mass Function

To effectively communicate the structure of a discrete probability distribution, we typically visualize probability mass functions using [bar charts](#) (or sometimes called probability histograms). Bar charts are the definitive standard for discrete distributions precisely because they graphically enforce the separation between each distinct, countable outcome and its corresponding probability mass. Unlike continuous histograms, which use contiguous bars, the gaps between the bars in a PMF visualization emphasize that probabilities exist only at specific points, not in the space between them.

The bar chart presented below illustrates the probabilities calculated for the Poisson distribution example discussed previously--specifically, the likelihood of various numbers of births occurring per hour. This visualization tool makes it immediately clear which outcomes are most likely (the modes) and how the probability quickly diminishes for rare events.



It is important to acknowledge that, mathematically, the possible number of events in a Poisson distribution (such as births per hour) extends theoretically to infinity. However, as demonstrated by the decreasing height of the bars, the probabilities assigned to outcomes far exceeding the mean ( $\lambda=2$ ) become rapidly and exponentially negligible. Consequently, outcomes greater than a certain threshold (like 10 in this visualization) often possess probabilities so low that they are effectively invisible on a standard bar chart.

## The Three Essential Axioms of a Valid PMF

For any mathematical function  $P(x)$  to be correctly identified and utilized as a valid **Probability Mass Function** in the field of **statistics**, it must flawlessly satisfy three fundamental properties derived from the broader axioms of probability theory. These rules ensure that the function accurately represents a real-world likelihood distribution:

**Non-Negativity:** The probability assigned to every outcome must be greater than or equal to zero. Mathematically,  $P(x) \geq 0$  for all  $x$ . This ensures that no outcome can have a negative likelihood, with probabilities being positive within the defined **support** (e.g.,  $P(X=1) = 1/6$ ) and exactly zero outside of it (e.g.,  $P(X=7) = 0$ ).

**Upper Bound:** The probability of any single, specific outcome must never exceed one.  $0 \leq P(x) \leq 1$ . This is a necessary condition, as it is impossible for a single event in any probability experiment to occur with a likelihood greater than 100%.

**Normalization:** The sum of the probabilities for every single possible outcome across the entire

**support** must aggregate precisely to one.  $\sum_{x} P(x) = 1$ . This final requirement confirms that the PMF completely defines the entire sample space, accounting for 100% of all potential results from the discrete random experiment.