

What is probability distribution?

A **probability distribution** is a function that describes all the possible values and outcomes for a random variable and the probability of each occurring. It's a way of representing the likelihood of different events.

Key Components

A probability distribution is fully defined by:

- **Sample Space:** The set of all possible outcomes. For a coin flip, the sample space is {Heads, Tails}.
- **Events:** Specific outcomes or a set of outcomes. For a six-sided die, "rolling a 4" is an event, as is "rolling an even number."
- **Probabilities:** The likelihood associated with each event.

There are two main types of probability distributions, depending on the nature of the data.

Types of Probability Distributions

1. Discrete Probability Distributions

This type describes the probabilities for a random variable that can only take on a finite or countably infinite number of values. The classic examples are a coin flip or a dice roll.

- **Example:** For a fair six-sided die, the probability of rolling any specific number (1, 2, 3, 4, 5, or 6) is exactly $1/6$.

2. Continuous Probability Distributions

This type describes the probabilities for a random variable that can take on any value within a continuous range. Examples include height, weight, or temperature. Instead of assigning a probability to a single point, it assigns a probability to a range of values.

Example: Rolling a Pair of Dice 🎲

Let's use a simple example: rolling two standard six-sided dice. The random event here is the sum of the numbers on the two dice.

- **Possible Outcomes:** The smallest sum is $1+1=2$, and the largest is $6+6=12$. So, the possible outcomes are all the integers from 2 to 12.
- **Probability of Each Outcome:** To find the probability, we can count the number of ways to get each sum and divide by the total number of possible combinations ($6 \times 6 = 36$).

Sum	Combinations (Die 1, Die 2)	Count	Probability (Count/36)
2	(1, 1)	1	$1/36 \approx 2.8\%$
3	(1, 2), (2, 1)	2	$2/36 \approx 5.6\%$
4	(1, 3), (2, 2), (3, 1)	3	$3/36 \approx 8.3\%$
5	(1, 4), (2, 3), (3, 2), (4, 1)	4	$4/36 \approx 11.1\%$
6	(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)	5	$5/36 \approx 13.9\%$
7	(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)	6	$6/36 \approx 16.7\%$
8	(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)	5	$5/36 \approx 13.9\%$
9	(3, 6), (4, 5), (5, 4), (6, 3)	4	$4/36 \approx 11.1\%$
10	(4, 6), (5, 5), (6, 4)	3	$3/36 \approx 8.3\%$
11	(5, 6), (6, 5)	2	$2/36 \approx 5.6\%$
12	(6, 6)	1	$1/36 \approx 2.8\%$

This table is the **probability distribution** for the sum of two dice. It shows that while all sums are possible, a sum of 7 is the most likely outcome, and a sum of 2 or 12 is the least likely.

Visualization

The probability distribution can also be visualized as a bar chart. This chart clearly shows the probability of each outcome, with the tallest bar representing

the most likely outcome (a sum of 7) and the shortest bars representing the least likely outcomes (sums of 2 and 12).