

What is a Categorical Distribution?

Authored by
stats writer

April 24, 2024

RECOMMENDED CITATION

stats writer (2024). *What is a Categorical Distribution?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=138757>

A categorical distribution is a statistical probability distribution that is used to represent the probability of outcomes that fall into a certain category. It is often used to model discrete data and is characterized by a finite set of categories, each with a corresponding probability. This type of distribution is commonly used in various fields such as market research, social sciences, and genetics to analyze and understand the frequency of events or outcomes within a given set of categories. In simple terms, a categorical distribution helps to determine the likelihood of different outcomes occurring within a specific set of options.

What is a Categorical Distribution?

A categorical distribution is a discrete probability distribution that describes the probability that a will take on a value that belongs to one of K categories, where each category has a probability associated with it.

For a distribution to be classified as a categorical distribution, it must meet the following criteria:

The categories are discrete. There are two or more potential categories. The probability that the random variable takes on a value in each category must be between 0 and 1. The sum of the probabilities for all categories must sum to 1.

The most obvious example of a categorical distribution is the distribution of outcomes associated with rolling a dice. There are $K = 6$ potential outcomes and the

probability for each outcome is $1/6$:

Example of a Categorical Distribution:

Outcomes of rolling a dice

Value	Probability
1	$1/6$
2	$1/6$
3	$1/6$
4	$1/6$
5	$1/6$
6	$1/6$

This distribution satisfies all of the criteria to be classified as a categorical distribution:

The categories are discrete (e.g. the random variable can only take on discrete values - 1, 2, 3, 4, 5, 6) There are two or more potential categories. The probability of each category is between 0 and 1. The sum of the probabilities add up to 1: $1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1$.

Rule of Thumb:

If you can count the number of outcomes, then you are working with a discrete random variable - e.g. counting

the number of times a coin lands on heads.

But if you can measure the outcome, you are working with a continuous random variable - e.g. measuring height, weight, time, etc.

Other Examples of Categorical Distributions

There are plenty of categorical distributions in the real world, including:

Example 1: Flipping a Coin.

When we flip a coin there are 2 potential discrete outcomes, the probability of each outcome is between 0 and 1, and the sum of the probabilities is equal to 1:

Example of a Categorical Distribution:

Outcomes of flipping a coin

Value	Probability
Heads	1/2
Tails	1/2

Example 2: Selecting Marbles from an Urn.

Suppose an urn contains 5 red marbles, 3 green

marbles, and 2 purple marbles. If we randomly select one marble from the urn, there are 3 potential discrete outcomes, the probability of each outcome is between 0 and 1, and the sum of the probabilities is equal to 1:

Example of a Categorical Distribution:

Outcomes of selecting a marble

Value	Probability
Red	5/10
Green	3/10
Purple	2/10

Example 3: Selecting a Card from a Deck.

Example of a Categorical Distribution:

Outcomes of selecting a card

Value	Probability
Ace	1/13
2	1/13
3	1/13
4	1/13
5	1/13
6	1/13
7	1/13
8	1/13
9	1/13
10	1/13
Jack	1/13
Queen	1/13
King	1/13

Relation to Other Distributions

For a distribution to be classified as a categorical distribution, it must have $K \geq 2$ potential outcomes and $n = 1$ trial.

Using this terminology, a categorical distribution is similar to the following distributions:

Bernoulli distribution: $K = 2$ outcomes, $n = 1$ trial

Binomial distribution: $K = 2$ outcomes, $n \geq 1$ trial

Multinomial distribution: $K \geq 2$ outcomes, $n \geq \text{trial}$

ARABPSYCHOLOGY.COM