

When should population standard deviation be used and when should sample standard deviation be used?

Authored by
stats writer

May 6, 2024

RECOMMENDED CITATION

stats writer (2024). *When should population standard deviation be used and when should sample standard deviation be used?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=143237>

The population standard deviation is a measure of the spread or variability of a data set that includes all members of a population. It is typically used when the entire population is available for analysis or when the goal is to make inferences about the entire population.

On the other hand, the sample standard deviation is a measure of the spread or variability of a sample data set. It is used when only a portion of the population is available for analysis or when the goal is to make inferences about the larger population based on the sample.

In general, the population standard deviation should be used when the data set includes all members of the population, while the sample standard deviation should be used when the data set only includes a portion of the population. Both measures serve different purposes and should be chosen based on the specific analysis being conducted.

Population vs. Sample Standard Deviation: When to Use Each

Standard deviation is one of the most common ways to measure the in a dataset.

It turns out that there are two different types of standard deviations you can calculate, depending on the type of data you're working with.

1. Population standard deviation

You should calculate the population standard deviation when the dataset you're working with represents an entire population, i.e. every value that you're interested in.

The formula to calculate a population standard deviation, denoted as σ , is:

$$\sigma = \sqrt{\sum(x_i - \mu)^2 / N}$$

where:

Σ : A symbol that means "sum"
 x_i : The i th value in a dataset
 μ : The population mean
 N : The population size

2. Sample standard deviation

You should calculate the sample standard deviation when the dataset you're working with represents a sample taken from a larger population of interest.

The formula to calculate a sample standard deviation, denoted as s , is:

$$s = \sqrt{\sum(x_i - \bar{x})^2 / (n - 1)}$$

where:

Σ : A symbol that means "sum"
 x_i : The i th value in a dataset
 \bar{x} : The sample mean
 n : The sample size

Population vs. Sample Standard Deviation: The Difference

From the formulas above, we can see that there is one tiny difference between the population and the sample standard deviation: When calculating the sample standard deviation, we divided by $n-1$ instead of N .

The reason for this is because when we calculate the sample standard deviation, we tend to underestimate the true variability in the population. In other words, our estimate of the true population standard deviation is biased.*

*Proof of this is beyond the scope of this article. For a mathematical proof, refer to from Stack Exchange.

Population vs. Sample Standard Deviation: When to Use Each

Use the following practice problems to gain a better understanding of when you should use population vs sample standard deviation.

Practice Problem 1: Sports

Suppose a basketball coach wants to summarize the mean and standard deviation of points scored by the 12 players on his team.

When calculating the standard deviation of points scored, should he use the population or sample standard deviation formula?

Answer: He should use the population standard deviation because he is only interested in the points scored by his players and not any other players on any other team.

Practice Problem 2: Height

Suppose a gym teacher wants to summarize the mean and standard deviation of heights of students in his class.

When calculating the standard deviation of height, should he use the population or sample standard deviation formula?

Answer: He should use the population standard deviation because he is only interested in the height of students in this one particular class.

Practice Problem 3: Biology

Suppose a biologist wants to summarize the mean and

standard deviation of the weight of a particular species of turtles. She decides to go out and collect a simple random sample of 20 turtles from the population.

When calculating the standard deviation of weights, should she use the population or sample standard deviation formula?

Answer: She should use the sample standard deviation because she is interested in the weights of the entire population of turtles, not just the weights of the turtles in her sample.

Practice Problem 4: Manufacturing

Suppose an inspector wants to summarize the mean and standard deviation of the weight of tires produced at a certain factory. He decides to collect a simple random sample of 40 tires from the factory and weighs each of them.

When calculating the standard deviation of weights, should he use the population or sample standard deviation formula?

Answer: He should use the sample standard deviation

because he is interested in the weights of all tires produced at this factory, not just the weights of the tires in his sample.

The following tutorials provide additional information about the standard deviation:

ARABPSYCHOLOGY.COM