Why is sample size important and can you provide some examples?

Authored by

stats writer

May 3, 2024

Sample size refers to the number of individuals or data points included in a study or experiment. It is an essential aspect of research as it directly affects the accuracy and reliability of the results. A larger sample size allows for more representative and generalizable findings, whereas a smaller sample size may lead to biased or misleading conclusions. For example, in a medical study evaluating the effectiveness of a new drug, a larger sample size would provide more reliable evidence compared to a smaller one. Similarly, in market research, a larger sample size would provide a more accurate understanding of consumer behavior and preferences. Overall, a sufficient sample size is crucial in ensuring the validity and credibility of research findings.

Why is Sample Size Important? (Explanation & Examples)

Sample size refers to the total number of individuals involved in an experiment or study.

Sample size is important because it directly affects how precisely we can estimate population parameters.

To understand why this is the case, it helps to have a basic understanding of confidence intervals.

A Brief Explanation of Confidence Intervals

In statistics, we're often interested in measuring - numbers that describe some characteristic of an entire population.

For example, we might be interested in measuring the mean height of all individuals in a certain city.

However, it's often too costly and time-consuming to go around and collect data on every individual in a population so we typically take a from the population instead and use data from the sample to estimate the population parameter.

For example, we might collect data on the height of 100 random individuals in the city. We can then calculate the mean height of the individuals in the sample. However, we can't be certain that the sample mean exactly matches the population mean.

To account for this uncertainty, we can create a . A confidence interval is a range of values that is likely to contain a population parameter with a certain level of confidence.

The formula to calculate a confidence interval for a population mean is:

Confidence Interval = x +/- z*(s/√n)

where:

x: sample meanz: the chosen z-values: sample standard deviationn: sample size

The z-value that you will use is dependent on the confidence level that you choose. The following table shows the z-value that corresponds to popular confidence level choices:

Confidence Level	z-value
0.90	1.645
0.95	1.96
0.99	2.58

The Relationship Between Sample Size & Confidence Intervals

Suppose we want to estimate the mean weight of a population of turtles. We collect a random sample of turtles with the following information:

Sample size n = 25Sample mean weight x = 300Sample standard deviation s = 18.5

Here is how to find calculate the 90% confidence interval for the true population mean weight:

90% Confidence Interval: 300 +/- 1.645*(18.5/√25) =

We are 90% confident that the true mean weight of the turtles in the population is between 293.91 and 306.09 pounds.

Now suppose instead of 25 turtles, we actually collect data for 50 turtles.

Here is how to find calculate the 90% confidence interval for the true population mean weight:

90% Confidence Interval: 300 +/- 1.645*(18.5/√50) =

Notice that this confidence interval is narrower than the previous confidence interval. This means our estimate of the true population mean weight of turtles is more precise.

Now suppose we instead collected data for 100 turtles.

Here is how to find calculate the 90% confidence interval for the true population mean weight:

90% Confidence Interval: 300 +/- 1.645*(18.5/√100) =

Notice that this confidence interval is *even narrower* than the previous confidence interval.

The following table summarizes each of the confidence interval widths:

Sample Size	90% Confidence Interval Width
25	12.18
50	8.51
100	6.08

Here's the takeaway: The larger the sample size, the more precisely we can estimate a population parameter.

The following tutorials provide other helpful explanations of confidence intervals and sample size.