## Estimation

Estimation in statistics are any procedures used to calculate the value of a population drawn from observations within a sample size drawn from that population. There are two types of estimation: either point or interval estimation.

## What are estimators?

Estimators are a concept in statistics that helps statisticians and researchers estimate the mean of a random sample. When the mean of a sample has statistical significance, the statistician can reasonably assume that it's an accurate reflection of a population's mean. This concept applies in scientific research, where the scientist tests a random sample of a larger population and generalizes their results to the total population. When scientists calculate the mean result of a sample and determine that the result has statistical significance, it means the results aren't likely to come from random chance.

**For example,** a pharmacologist testing a new migraine medication selects a random sample of 100 participants to try the medication. The participants self-report how frequently they experience migraine symptoms over the course of a month. Then, the researcher calculates the mean number of days that the sample reports having migraine symptoms. They find the sample experienced a mean of 10 days of symptoms. This number represents the sample mean, also known as an estimator, and allows the researcher to estimate that a more general population is likely to experience 10 days of migraine symptoms per month while taking the medication.

## Types of estimators

The two types of estimators are point and interval estimators. Here are definitions for both types of estimators:

## Point estimator

A point estimator allows the researcher or statistician to calculate a single value estimate of a parameter. In statistics, a parameter is a characteristic of a population which the researcher or statistician wants to assess. Examples of parameters include the population mean, the population variance and the population proportion. Scientists use these characteristics to find sample statistics, which include the sample mean, variance and proportion. They use point estimators when finding a single value estimate of a parameter. For example, finding the mean of a sample is a single value that represents a population mean.

A researcher might use a point estimator to estimate the proportion of elementary school students who ride the bus to school. In this example, the researcher plans to estimate what proportion of the total population of students at a school rides the bus by surveying the transportation methods of a random sample of students. They find the ratio of bus riders to non-bus riders within their sample, and they generalize this percentage to the total population. For example, they may find that 50% of their sample rides the bus, meaning that approximately 50% of all students ride the bus.

## Interval estimator

An interval estimator assesses the interval of possible values that an unknown population parameter can represent. Researchers use this type of estimator to calculate a range of possible values for a given parameter and determine the likelihood of each being true. They assess the level of possibility for each value in the estimate range by calculating its confidence interval. A confidence interval represents the likelihood of a finding being true. Calculating this function allows scientists to determine the chance that each value in a range of possible values representing the parameter is likely to be accurate.

**For example,** a researcher might use an interval estimator to calculate the reliability of data collected in a survey. A survey of customers for an online retailer shows that 75% of respondents plan to purchase from the company again. The researcher conducts a study to collect data about how many customers make repeat purchases from the company. They test whether the proportion of repeat purchasers is equal to the 75% of survey takers who expressed interest in purchasing from the retailer again. If the proportion of repeat purchasers closely matches the data from the survey, then the survey is accurate.

## Point estimation vs. interval estimation

Here are the key differences between point and interval estimation with considerations for when to use each:

### Values tested

Point and interval estimates test different types of values. While point estimates focus on testing single values, such as the mean of a characteristic represented within a sample, interval estimates assess a range of values. For example, a researcher can use a point estimate to find the mean age of a sample and use it to predict the average age of those within a population. Age represents a single value in this example.

By comparison, the researcher may use an interval estimate to calculate the proportion of the sample over the age of 32 who experience migraine symptoms for ten or more days each month. Finding this data requires multiple values. The researcher collects data about the ages of participants within their sample and the number of days they experience migraine symptoms each month.

### Applications

Since point and interval estimators measure different values, they have different applications. Typically, researchers use point estimators to find sample means. This is a single value that they use to generalize to the total population. Researchers use interval estimators to assess the accuracy or credibility of different values in a range.

### Interpretations

A point estimate offers the best possible estimate for the mean of a population parameter. Although not always accurate, a point estimate often provides a basis on which a researcher can interpret a statistic. This estimate assumes that the average of a sufficiently larger sample

population is a reasonable reflection of the mean parameter of a population, allowing the researcher to conduct further statistical analyses with this assumption as a foundation.

Interval estimators are important for helping researchers understand the degree of uncertainty in the estimate of a parameter. This uncertainty level allows researchers to consider diverse factors that may contribute to differences in the confidence level between possible variables. Based on their interpretation, they may conduct additional studies or analyses to better understand the mechanisms that cause differences between values.

## Descriptive terms for estimators

Here are some common terms used to describe types of estimators:

**Biased:** Bias refers to a statistic that either underestimates or overestimates a value.

**Efficient:** An efficient statistic is one with minimal variance. Variance is the measure of how significantly a set of numbers spreads out from the average value.

**Invariant:** Invariant statistics are values that don't change easily as the result of a transformation or data shift.

**Shrinkage**: Shrinkage is a piece of raw data that researchers can improve by combining with other data.

**Sufficient:** Sufficient statistics summarize all the available data about a sample within a parameter.

**Unbiased:** An unbiased statistic is one that neither overestimates nor underestimates a value, meaning it's an accurate reflection of a given value.