How can a Trimmed Mean be calculated in R, and can you provide some examples?

Authored by

stats writer

May 6, 2024

A Trimmed Mean is a statistical measure used to calculate the average of a dataset while removing a certain percentage of extreme values from the dataset. This can be done in R by using the "trim" function, which allows you to specify the percentage of values to be trimmed. For example, if you have a dataset of 100 values, and you want to calculate the trimmed mean by removing the top and bottom 10 values, you would use the "trim" function with the argument "trim=10". This will calculate the mean of the remaining 80 values. Another example could be calculating the trimmed mean by removing the top and bottom 25% of values, using the argument "trim=25". This function can be useful in situations where extreme values may skew the overall mean of the dataset.

Calculate a Trimmed Mean in R (With Examples)

A trimmed mean is the mean of a dataset that has been calculated after removing a specific percentage of the smallest and largest values from the dataset.

For example, a 10% trimmed mean would represent the mean of a dataset after the 10% smallest values and 10% largest values have been removed.

The easiest way to calculate a trimmed mean in R is to use the following basic syntax:

```
#calculate 10% trimmed mean
mean(x, trim=0.1)
```

The following examples show how to use this function to calculate a trimmed mean in practice.

Example 1: Calculate Trimmed Mean of Vector

The following code shows how to calculate a 10% trimmed mean for a vector of data:

```
#define data
data = c(22, 25, 29, 11, 14, 18, 13, 13, 17, 11, 8, 8, 7, 12, 15, 6, 8, 7, 9, 12)

#calculate 10% trimmed mean
mean(data, trim=0.1)

12.375
```

The 10% trimmed mean is 12.375.

This is the mean of the dataset after the smallest 10% and largest 10% of values have been removed from the dataset.

Example 2: Calculate Trimmed Mean of Column in Data Frame

The following code shows how to calculate a 5% trimmed mean for a specific column in a data frame:

```
#create data frame
df = data.frame(points=c(25, 12, 15, 14, 19, 23, 25, 29),
```

```
assists=c(5, 7, 7, 9, 12, 9, 9, 4),
rebounds=c(11, 8, 10, 6, 6, 5, 9, 12))

#calculate 5% trimmed mean of points
mean(df$points, trim=0.05)
```

```
20.25
```

The 5% trimmed mean of the values in the 'points' column is 20.25.

This is the mean of the 'points' column after the smallest 5% and largest 5% of values have been removed.

Example 3: Calculate Trimmed Mean of Multiple Columns

The following code shows how to calculate a 5% trimmed mean for multiple columns in a data frame:

```
#create data frame
df = data.frame(points=c(25, 12, 15, 14, 19, 23, 25, 29),
assists=c(5, 7, 7, 9, 12, 9, 9, 4),
rebounds=c(11, 8, 10, 6, 6, 5, 9, 12))

#calculate 5% trimmed mean of points and assists
```

sapply(df, function(x) mean(x, trim=0.05))

points assists
20.25 7.75

From the output we can see:

The 5% trimmed mean of the 'points' column is 20.25.The 5% trimmed mean of the 'assists' column is 7.75.

The following tutorials provide additional information about trimmed means: