

M2.4: Statistical Analysis

Mean

Statistical analysis in R is performed by using many in-built functions. Most of these functions are part of the R base package. These functions take R vector as an input along with the arguments and give the result.

The functions we are discussing in this chapter are mean, median and mode.

Mean

It is calculated by taking the sum of the values and dividing with the number of values in a data series.

The function **mean()** is used to calculate this in R.

Syntax

The basic syntax for calculating mean in R is –

```
mean(x, trim = 0, na.rm = FALSE, ...)
```

Following is the description of the parameters used –

- **x** is the input vector.
- **trim** is used to drop some observations from both end of the sorted vector.
- **na.rm** is used to remove the missing values from the input vector.

```
# Create a vector.  
x <- c(12,7,3,4.2,18,2,54,-21,8,-5)  
  
# Find Mean.  
result.mean <- mean(x)  
print(result.mean)
```

When we execute the above code, it produces the following result –

```
[1] 8.22
```

Applying Trim Option

M2.4: Statistical Analysis

When trim parameter is supplied, the values in the vector get sorted and then the required numbers of observations are dropped from calculating the mean.

When trim = 0.3, 3 values from each end will be dropped from the calculations to find mean.

In this case the sorted vector is (-21, -5, 2, 3, 4.2, 7, 8, 12, 18, 54) and the values removed from the vector for calculating mean are (-21,-5,2) from left and (12,18,54) from right.

```
# Create a vector.  
x <- c(12,7,3,4.2,18,2,54,-21,8,-5)  
  
# Find Mean.  
result.mean <- mean(x,trim = 0.3)  
print(result.mean)
```

When we execute the above code, it produces the following result –

```
[1] 5.55
```

Applying NA Option

If there are missing values, then the mean function returns NA.

To drop the missing values from the calculation use na.rm = TRUE. which means remove the NA values.

```
# Create a vector.  
x <- c(12,7,3,4.2,18,2,54,-21,8,-5,NA)  
  
# Find mean.  
result.mean <- mean(x)  
print(result.mean)  
  
# Find mean dropping NA values.  
result.mean <- mean(x,na.rm = TRUE)  
print(result.mean)
```

Unique solution ID: #1092

Author: UserAdmin

Last update: 2020-01-09 09:57