

How can I calculate the pooled variance in R?

Authored by
stats writer

April 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate the pooled variance in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=140154>

Calculating the pooled variance in R involves combining the variances from two or more samples to obtain a single measure of variance that takes into account the size and variability of each sample. This can be achieved by using the "pool" function, which takes in the individual variances and sample sizes as inputs and outputs the pooled variance. This method is commonly used in statistical analysis to determine the overall variance of a population when multiple samples are available. By utilizing the "pool" function in R, researchers can efficiently and accurately calculate the pooled variance, providing a more comprehensive understanding of the data and its variability.

Calculate Pooled Variance in R

In statistics, refers to the average of two or more group variances.

We use the word "pooled" to indicate that we're "pooling" two or more group variances to come up with a single number for the common variance between the groups.

In practice, pooled variance is used most often in a , which is used to determine whether or not two population means are equal.

The pooled variance between two samples is typically denoted as sp^2 and is calculated as:

$$sp^2 = ((n1-1)s1^2 + (n2-1)s2^2) / (n1+n2-2)$$

Unfortunately there is no built-in function to calculate

the pooled variance between two groups in R, but we can calculate it fairly easily.

For example, suppose we want to calculate the pooled variance between the following two groups:

Dataset 1	Dataset 2
6	5
7	7
7	7
8	8
10	10
11	13
13	14
14	15
14	19
16	20
18	20
19	23
19	25
19	28
20	32

The following code shows how to calculate the pooled variance between these groups in R:

```
#define groups of data
```

```
x1 <- c(6, 7, 7, 8, 10, 11, 13, 14, 14, 16, 18, 19, 19, 19, 20)
```

```
x2 <- c(5, 7, 7, 8, 10, 13, 14, 15, 19, 20, 20, 23, 25, 28, 32)
```

```
#calculate sample size of each group
```

```
n1 <- length(x1)
n2 <- length(x2)

#calculate sample variance of each group
var1 <- var(x1)
var2 <- var(x2)

#calculate pooled variance between the two groups
pooled <- ((n1-1)*var1 + (n2-1)*var2) / (n1+n2-2)

#display pooled variance
pooled

46.97143
```

The pooled variance between these two groups turns out to be 46.97143.