Standardabweichung berechnen

Table of contents

Standard deviation (σ) equals the squareroot ($\sqrt{}$) of the squared (^{2}) sum ($\sum$) of the difference between each value and the mean ($x_{1...n}-\mu$), divided by the number of values minus 1 ($N-1$):

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N - 1}} \tag{1}$$

Let's try doing this step-by-step:

1. Save values as object
2. Determine N
3. compute mean (μ)
4. compute difference between each value and the mean (subtract mean from each value; $x_{1...n} - \mu$)

5. square these values (^{2})
6. sum the squared values ($\sum$)
7. divide this value by N-1
8. compute the squareroot ($\sqrt{\ }$)

Save values as object

```
werte <- c(13, 0, 2000, 5, 19)
```

Print values

```
werte
```

```
[1]   13    0 2000    5   19
```

Determine N

* By hand: print values and count how many there are

```
werte
```

```
[1]   13    0 2000    5   19
```

* Or, use the `length()` function

```
length(werte)
```

```
[1] 5
```

Compute mean

The mean (μ) is the sum ($\sum$) of the values (x) divided by the number of values (N):

$$\mu = \frac{\sum x}{N}$$

* By hand

Typing out all the values:

```
(13 + 0 + 2000 + 5 + 19)/5
```

[1] 407.4

* Or by using the functions `sum()` and `length()`

```
sum(werte)/length(werte)
```

[1] 407.4

* Or by simply using the function `mean()`

```
mean(werte)
```

[1] 407.4

Subtract mean from each value

This can also be done several different ways

* By hand (tedious and prone to human error)

```
(13 - 407.4)
```

[1] -394.4

```
(0 - 407.4)
```

[1] -407.4

```
(2000 - 407.4)
```

[1] 1592.6

```
(5 - 407.4)
```

```
[1] -402.4
```

```
(19 - 407.4)
```

```
[1] -388.4
```

* Or using the object name

```
werte - mean(werte)
```

```
[1] -394.4 -407.4 1592.6 -402.4 -388.4
```

Square these values

* By hand

```
(13 - 407.4)^2
```

```
[1] 155551.4
```

```
(0 - 407.4)^2
```

```
[1] 165974.8
```

```
(2000 - 407.4)^2
```

```
[1] 2536375
```

```
(5 - 407.4)^2
```

```
[1] 161925.8
```

```r
(19 - 407.4)^2
```

```
[1] 150854.6
```

* Or by using ^2 with the variable name

```r
(werte - mean(werte))^2
```

```
[1]  155551.4  165974.8 2536374.8  161925.8  150854.6
```

Sum the squared values

* By hand

```r
(13 - 407.4)^2 +
(0 - 407.4)^2 +
(2000 - 407.4)^2 +
(5 - 407.4)^2 +
(19 - 407.4)^2
```

```
[1] 3170681
```

* Or with the object name and some functions

```r
sum((werte - mean(werte)) ^ 2)
```

```
[1] 3170681
```

Divide by N-1

* By hand

```r
((13 - 407.4) ^ 2 +
   (0 - 407.4) ^ 2 +
   (2000 - 407.4) ^ 2 +
```

```
    (5 - 407.4) ^ 2 +
    (19 - 407.4) ^ 2)/(5-1)
```

[1] 792670.3

* Or with the object name and some functions

```
sum((werte - mean(werte)) ^ 2) / (length(werte) - 1)
```

[1] 792670.3

Compute the squareroot

* By hand

```
sqrt(((13 - 407.4)^2 +
(0 - 407.4)^2 +
(2000 - 407.4)^2 +
(5 - 407.4)^2 +
(19 - 407.4)^2)/(5-1))
```

[1] 890.3203

* Or with the object name and some functions

```
sqrt(sum((werte - mean(werte)) ^ 2) / (length(werte) - 1))
```

[1] 890.3203

Check with `sd()` function

```
sd(werte)
```

[1] 890.3203