



Statistical Variance

Siddharth Kalla, Lyndsay T Wilson 178.7K reads

Statistical variance gives a measure of how the data distributes itself about the mean or expected value. Unlike range that only looks at the extremes, the variance looks at all the data points and then determines their distribution.

In many cases of statistics and experimentation, it is the variance that gives invaluable information about the data distribution.

EXPLORABLE
Quiz Time!

Quiz: Psychology 101 Part 2

Quiz: Psychology 101 Part 2

Quiz: Flags in Europe

[See all quizzes =>](#)

Variance Calculation (population of Scores)

The mathematical formula to calculate the variance is given by:

$$\sigma^2 = \frac{\sum(X - \mu)^2}{N}$$

σ^2 = variance

? $(X - \mu)^2$ = The sum of $(X - \mu)^2$ for all datapoints

X = individual data points

μ = mean [1] of the population

N = number of data points

This means the square of the variance is given by the average of the squares of difference between the data points and the mean.

Step By Step Calculation

For example, suppose you want to find the variance of scores on a test. **Suppose the scores are 67, 72, 85, 93 and 98.**

First, write down the formula for variance:

$$s^2 = \frac{\sum (x - \mu)^2}{N}$$

Next, there are five scores in total, so $N = 5$.

$$s^2 = \frac{\sum (x - \mu)^2}{5}$$

Calculate the mean (μ) for the five scores: $67 + 72 + 85 + 93 + 98 / 5$, so $\mu = 83$.

$$s^2 = \frac{\sum (x - 83)^2}{5}$$

Now, **compare each score** ($x = 67, 72, 85, 93, 98$) **to the mean** ($\mu = 83$)

$$s^2 = \frac{[(67 - 83)^2 + (72 - 83)^2 + (85 - 83)^2 + (93 - 83)^2 + (98 - 83)^2]}{5}$$

Conduct the subtraction in each paranthesis.

$$67 - 83 = -16$$

$$72 - 83 = -11$$

$$85 - 83 = 2$$

$$93 - 83 = 10$$

$$98 - 83 = 15$$

The formula will now look like this:

$$s^2 = \frac{[(-16)^2 + (-11)^2 + (2)^2 + (10)^2 + (15)^2]}{5}$$

Then, **square each paranthesis.** We get 256, 121, 4, 100 and 225.

This is how:

$$s^2 = [(-16) \times (-16) + (-11) \times (-11) + (2) \times (2) + (10) \times (10) + (15) \times (15)] / 5$$

$$s^2 = [16 \times 16 + 11 \times 11 + 2 \times 2 + 10 \times 10 + 15 \times 15] / 5$$

which equals:

$$s^2 = [256 + 121 + 4 + 100 + 225] / 5$$

7 Then summarize the numbers inside the brackets:

$$s^2 = 706 / 5$$

To get the final answer, we **divide the sum by 5** (Because it was a total of five scores). This is the final variance for the dataset:

$$s^2 = \underline{\underline{141.2}}$$

This is the variance of the population of scores.

Variance of a Sample

In many cases, instead of a population, we deal with samples.

In this case, we need to slightly change the formula for variance to:

$$s^2 = \frac{\sum (X - \bar{X})^2}{n - 1}$$

s^2 = the variance of the sample.

Note that the denominator is one less than the sample size in this case.

Usage

The concept of variance can be extended to continuous data sets too. In that case, instead of summing up the individual differences from the mean, we need to integrate them. This approach is also useful when the number of data points is very large, for example the population of a country.

Variance is extensively used in probability theory, where from a given smaller sample set, more generalized conclusions need to be drawn. This is because variance gives us an idea about the distribution of data around the mean, and thus from this distribution, we can work out where we can expect an unknown data point.

Links

[1] <https://explorable.com/arithmetric-mean>