

# Variance

Authored by  
**mohammad looti**

October 8, 2025

## RECOMMENDED CITATION

mohammad looti (2025). *Variance*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=36290>

## Variance

**Primary Disciplinary Field(s):** Statistics, Probability Theory, Data Science, Econometrics

### 1. Core Definition

Variance, denoted typically as  $\sigma^2$  (for population variance) or  $s^2$  (for sample variance), is a fundamental measure in descriptive statistics and probability theory. It quantifies the extent of the spread or dispersion of a set of numerical data points relative to their central tendency, specifically the **mean**. In essence, variance answers the question: Do the scores tend to cluster tightly around the mean, or are they broadly scattered? Mathematically, variance is defined as the expected value of the squared deviation from the mean, meaning that it measures the average squared distance of each point from the distribution's center. Because the deviations are squared, variance is always a non-negative value, and zero variance implies that all values within the dataset are identical.

The core utility of variance lies in its ability to differentiate between datasets that might possess identical means but vastly different characteristics regarding consistency or risk. The original source content provides a simple yet effective demonstration of this principle. Consider two data sets, both possessing seven scores ranging from 1 to 9, and both having a mean of 5. The first set, , exhibits scores that frequently deviate significantly toward the extremes, resulting in a large average squared distance from the mean. Conversely, the second set, , shows values clustering much closer to the center, leading to a much smaller overall variance. Thus, high variance signifies a wide dispersion and potential unpredictability in the data, while low variance suggests consistency and high reliability around the average measure.

Understanding the calculation involves three key steps: first, calculating the difference between each individual data point and the mean; second, squaring these differences to ensure positive values (and to amplify the influence of extreme outliers); and third, averaging these squared differences across the entire dataset. This squaring operation is crucial because, if the deviations were not squared, the sum of deviations from the mean would always equal zero, rendering the measure useless for determining spread. The resulting numerical value of the variance, however, exists in units squared relative to the original data units, which is a key characteristic that necessitates its complement, the **standard deviation**, for practical interpretation.

### 2. Mathematical Formulation and Calculation

The exact formulation of variance depends critically on whether one is analyzing an entire population or merely a sample drawn from that population. For a finite population of size  $N$ , where  $\mu$  is the population mean, the population variance ( $\sigma^2$ ) is calculated using the following formula:  $\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$ . This formula represents the

true, average squared deviation for every member of the complete group being studied. Calculating the population variance requires access to every data point, which is often infeasible in real-world statistical applications, particularly in fields like economics or physics where populations are theoretically infinite or practically too large to census entirely.

When working with a **sample** of size  $n$ , the sample variance ( $s^2$ ) must be computed. A direct application of the population formula (dividing by  $n$ ) would yield a result that consistently underestimates the true population variance, a phenomenon known as bias. To correct for this systematic bias and ensure that the sample variance is an **unbiased estimator** of the population variance, statisticians employ what is known as Bessel's correction. This correction involves dividing the sum of squared deviations not by  $n$ , but by  $n-1$ , corresponding to the degrees of freedom. Thus, the sample variance formula is  $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$ , where  $\bar{x}$  is the sample mean. The use of  $n-1$  accounts for the fact that one degree of freedom is 'used up' in calculating the sample mean, which is necessary before deviations can be measured.

In many computational contexts, especially when the mean is a non-integer or when dealing with large datasets, calculating the variance using the deviation method described above can be cumbersome and prone to rounding errors. An equivalent and often more computationally stable formula, derived from properties of expectation, is utilized:  $\sigma^2 = E - (E)^2$ . This alternative form states that the variance is equal to the expected value of the square of the variable minus the square of the expected value (mean) of the variable. This approach is frequently preferred in complex probability distributions and algorithmic calculations, streamlining the process by avoiding the intermediate step of calculating the deviations from the mean for every single data point before squaring and summing them.

### 3. Relationship to Standard Deviation and Interpretation

While variance is the fundamental measure of dispersion, it is often not the most intuitive measure for direct interpretation, primarily because its units are squared. If data points are measured in meters, the variance is expressed in square meters, which lacks practical meaning for describing distance. To resolve this issue, statisticians rely heavily on the **standard deviation** ( $\sigma$  or  $s$ ), which is simply the positive square root of the variance. By taking the square root, the measure is returned to the original units of the data, providing an easily interpretable measure of the typical deviation from the mean.

The standard deviation allows for direct contextual assessment. For instance, in a normal distribution, approximately 68% of data points fall within one standard deviation of the mean, and 95% fall within two standard deviations. This rule of thumb, derived from Chebyshev's inequality and the empirical rule for normal distributions, allows analysts to immediately gauge the

consistency of the data and identify potential outliers. Therefore, while variance provides the mathematical foundation and possesses superior additive properties (as discussed below), the standard deviation serves as the primary tool for communicating spread and variability in applied reports and analyses.

The choice between using variance and standard deviation often depends on the task at hand. In advanced mathematical statistics, particularly when dealing with the combination of random variables or within inferential statistics like the Analysis of Variance (ANOVA), variance is the preferred metric due to its linear property of additivity. When independent random variables are summed, their variances add directly. Conversely, standard deviations do not possess this simple additive property, making variance indispensable for theoretical modeling and complex statistical hypothesis testing.

#### 4. Historical Development and Theoretical Origin

Although the concept of measuring dispersion has roots dating back to the development of the least squares method by **Carl Friedrich Gauss** and Adrien-Marie Legendre in the early 19th century, the term "variance" and its precise theoretical application were formalized much later by the pioneering British statistician and geneticist, **Sir Ronald Fisher**. Fisher introduced the term in his 1918 paper, "The Correlation Between Relatives on the Supposition of Mendelian Inheritance." This work was crucial not only for mathematical genetics but also for establishing a rigorous framework for modern statistical inference.

Fisher's introduction of variance was intrinsically linked to his development of the Analysis of Variance (ANOVA). ANOVA provided a methodology for partitioning the total observed variability (total variance) in a dataset into different components attributable to different sources or factors (e.g., separating variance due to experimental treatment from variance due to measurement error). This ability to decompose variability revolutionized fields ranging from agricultural science (where Fisher initially applied these methods) to experimental psychology, establishing variance as the cornerstone measure for inferential statistics that seek to determine the causes of observed differences.

Prior to Fisher, alternative measures of dispersion were more common, notably the **mean absolute deviation (MAD)**, which averages the absolute (non-squared) differences from the mean. While mathematically simpler, MAD proved less amenable to algebraic manipulation and the theoretical development of normal distribution properties that Fisher and his contemporaries were pursuing. The squaring of deviations inherent in variance provides algebraic tractability, particularly in complex multivariate analyses, thereby cementing variance as the standard measure in classical statistics despite the interpretive simplicity of its competitors.

## 5. Key Properties of Variance

The utility of variance in probability theory and advanced statistics stems from several critical mathematical properties that define its behavior under transformations and combinations of random variables. Perhaps the most fundamental property is that variance is always non-negative:  $\text{Var}(X) \geq 0$ . Since variance is defined as the average of squared differences, it is mathematically impossible for a distribution to have a negative variance. A variance of exactly zero signifies a deterministic variable--one that never deviates from its mean value.

Another essential property governs how variance responds to linear transformations of the data. If  $X$  is a random variable and  $a$  and  $b$  are constants, the variance of the transformed variable  $Y = aX + b$  is given by  $\text{Var}(Y) = a^2 \text{Var}(X)$ . This reveals two important facts: first, adding a constant ( $b$ ) to every data point (a shift in location) does not affect the spread, and thus variance remains unchanged. Second, multiplying the variable by a scaling factor ( $a$ ) scales the variance by the square of that factor ( $a^2$ ). This squared relationship reinforces why the standard deviation, being  $\sqrt{\text{Var}(X)}$ , scales linearly by  $|a|$ , simplifying interpretation of spread when units are converted.

The most powerful property, foundational to many statistical tests, relates to the variance of the sum or difference of independent random variables. If  $X_1, X_2, \dots, X_k$  are mutually independent random variables, the variance of their sum or difference is simply the sum of their individual variances:  $\text{Var}(X_1 + X_2 + \dots + X_k) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_k)$ . This additive property holds even if the variables are subtracted (i.e.,  $\text{Var}(X_1 - X_2) = \text{Var}(X_1) + \text{Var}(X_2)$ ). This feature is central to proving theorems like the Central Limit Theorem and underlies the mathematical structure of complex statistical models, including regression analysis and portfolio diversification models, where risks (variances) are aggregated.

## 6. Applications Across Disciplines

Variance serves as a crucial metric across virtually all fields utilizing quantitative data, acting as a primary proxy for risk, quality, and consistency. In the financial sector, variance, or its square root, volatility, is the standard measure of investment risk. According to **Modern Portfolio Theory (MPT)**, rational investors seek to maximize returns for a given level of risk, where risk is strictly defined as the variance of portfolio returns. High variance in stock returns signifies high volatility and unpredictability, while low variance suggests a more stable, predictable asset.

In manufacturing and quality control, variance is used extensively to monitor process stability. A production process must aim for low variance in product specifications (e.g., thickness, weight, strength) to ensure consistency and meet engineering tolerances. Statistical Process Control (SPC) charts rely on variance estimates to set control limits; if the variance of a measured

characteristic increases beyond acceptable limits, it signals that the process is out of control and requires immediate intervention to identify and correct the source of the increased variability.

Beyond engineering and finance, variance is indispensable in experimental design in social sciences and biology. When researchers conduct controlled experiments, they use variance analysis (ANOVA) to determine if observed differences in outcomes between experimental groups are statistically significant or merely due to random chance (inherent variability within the population). By comparing the variance explained by the treatment (between-group variance) to the variance unexplained (within-group error variance), researchers can draw robust causal inferences about the effectiveness of interventions.

## 7. Criticisms and Limitations

Despite its ubiquity and mathematical elegance, variance possesses significant limitations, primarily stemming from the operation of squaring the deviations. This squaring process means that variance is highly sensitive to **outliers** or extreme values. A single data point that lies far from the mean can contribute disproportionately to the total variance, potentially skewing the perception of spread for the majority of the data. For datasets where extreme events are common but not representative of the typical spread, variance can be a misleading metric of dispersion.

A second key criticism relates to the lack of robustness. Unlike robust statistical measures, variance provides no meaningful resistance to measurement error or data contamination. Researchers in robust statistics often prefer alternatives such as the Interquartile Range (IQR) or the Mean Absolute Deviation (MAD). The IQR, which measures the spread of the middle 50% of the data, is entirely unaffected by the most extreme upper and lower quartiles, offering a stable measure of central dispersion regardless of outliers.

Furthermore, for distributions that are highly non-normal--such as those that are heavily skewed or possess extreme kurtosis (fat tails)--variance can be difficult to interpret within the context of standard statistical models. While variance is a sufficient descriptor for many Gaussian (normal) phenomena, its application to processes governed by, for example, power-law distributions (common in areas like network analysis or income distribution) may mask important distributional characteristics, leading to an incomplete or inappropriate risk assessment. In these cases, methods that focus on higher moments or non-parametric statistics are often deemed more appropriate than reliance solely on the variance and mean.

## Further Reading

[Variance \(Wikipedia\)](#)

[Standard Deviation](#)

[Ronald Fisher and the Development of ANOVA](#)

---

Bessel's Correction and Unbiased Estimators

ARABPSYCHOLOGY.COM