

How to Easily Understand the Difference Between Standardization and Normalization

Authored by
stats writer

December 4, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Understand the Difference Between Standardization and Normalization*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=105316>

In the world of statistics and machine learning, preparing raw data for analysis is a critical step often involving feature scaling. Two fundamental techniques used for this purpose are **standardization** and **normalization**. While both methods aim to transform numerical data into a comparable range, they operate under distinct principles and produce fundamentally different results.

Standardization involves converting a dataset so that all values possess a mean of 0 and a standard deviation of 1. In contrast, normalization rescales the values so that they fall within a fixed range, typically between 0 and 1. The key conceptual difference lies in their impact on the distribution: standardization preserves the original shape of the data's distribution, while normalization, being dependent on minimum and maximum values, can potentially compress or stretch the distribution shape during the transformation.

Understanding Data Rescaling in Statistics

Data scaling methods like **standardization** and **normalization** are essential preprocessing steps, particularly when working with algorithms that rely on distance calculations, such as K-Nearest Neighbors or Support Vector Machines. If features in a dataset are measured on dramatically different scales, the features with larger magnitudes might disproportionately influence the final model, leading to biased results.

To mitigate this issue, we employ scaling techniques. Although sometimes used interchangeably in casual conversation, these terms refer to precise statistical processes. Understanding the mathematical foundation and practical implications of each method is crucial for selecting the appropriate approach for your specific analytical goals.

We will now delve into the precise mathematical definitions and formulas governing **standardization** and **normalization**.

Defining Standardization (Z-Score Scaling)

Standardization, often referred to as Z-score scaling, transforms the dataset so that it follows a standard normal distribution. This results in a mean (average) of zero and a standard deviation of one. This technique is highly effective when the feature distribution is approximately Gaussian (bell-shaped) or when algorithms assume a zero-centered input.

The primary benefit of standardization is that it handles outliers gracefully, as the transformation does not bound the data to a small range. The standardized value (or Z-score) represents the number of standard deviations a data point is away from the mean.

The formula used for **Standardization** is:

$$x_{\text{new}} = (x_i - x) / s$$

where:

x_i : Represents the i th individual value in the dataset being transformed.

x : Denotes the sample arithmetic mean of the feature column.

s : Represents the sample standard deviation of the feature column.

Defining Normalization (Min-Max Scaling)

Normalization, specifically Min-Max scaling, adjusts the dataset values so that every observation falls within a fixed, predefined interval, almost always between 0 and 1. This scaling is particularly useful when the distribution of the data is unknown or when the data must adhere to strict boundary conditions (e.g., in image processing where pixel values range from 0 to 255, and often need to be mapped to 0 to 1).

A significant drawback of normalization is its susceptibility to outliers. Since the transformation relies heavily on the absolute minimum (x_{min}) and maximum (x_{max}) values, a single extreme outlier can compress the vast majority of the "normal" data points into a very narrow range, reducing their discriminatory power.

The formula used for **Normalization** (Min-Max Scaling) is:

$$x_{\text{new}} = (x_i - x_{\text{min}}) / (x_{\text{max}} - x_{\text{min}})$$

where:

x_i : The i th individual value in the dataset.

x_{min} : The absolute minimum value found in the feature column.

x_{max} : The absolute maximum value found in the feature column.

Practical Demonstration: How to Standardize a Dataset

To illustrate the process of standardization, let us consider a sample numerical dataset. We will demonstrate how each raw value is transformed into a corresponding Z-score.

Suppose we begin with the following raw data points:

Data Values
13
16
19
22
23
38
47
56
58
63
65
70
71

For this specific dataset, the calculated mean is 43.15, and the standard deviation is 22.13. We use these parameters to apply the standardization formula (Z-score scaling) to individual observations.

To standardize the first value, **13**, we apply the formula:

$$\mathbf{x_{new} = (x_i - \bar{x}) / s = (13 - 43.15) / 22.13 = -1.36}$$

Next, to standardize the second value, **16**, we utilize the same process:

$$\mathbf{x_{new} = (x_i - \bar{x}) / s = (16 - 43.15) / 22.13 = -1.23}$$

And finally, for the third value, **19**:

$$\mathbf{x_{new} = (x_i - \bar{x}) / s = (19 - 43.15) / 22.13 = -1.09}$$

By applying this formula to all data points, we achieve a standardized dataset where the new values represent the number of standard deviations away from the zero mean. The resulting dataset appears as follows:

Data Values	Standardized
13	-1.36
16	-1.23
19	-1.09
22	-0.96
23	-0.91
38	-0.23
47	0.17
56	0.58
58	0.67
63	0.90
65	0.99
70	1.21
71	1.26

Practical Demonstration: How to Normalize a Dataset

Let us use the same raw dataset to demonstrate **normalization** using the Min-Max scaling technique, ensuring all resultant values fall strictly between 0 and 1.

The initial dataset remains:

Data Values
13
16
19
22
23
38
47
56
58
63
65
70
71

To calculate the normalized values, we first identify the boundary points: the minimum value (x_{min}) in this set is 13, and the maximum value (x_{max}) is 71. The difference ($x_{max} - x_{min}$) is therefore 58.

To normalize the first value, **13**, we apply the Min-Max formula:

$$x_{new} = (x_i - x_{min}) / (x_{max} - x_{min}) = (13 - 13) / (71 - 13) = 0$$

To normalize the second value, **16**:

$$x_{new} = (x_i - x_{min}) / (x_{max} - x_{min}) = (16 - 13) / (71 - 13) = 3 / 58 \approx .0517$$

To normalize the third value, **19**:

$$x_{new} = (x_i - x_{min}) / (x_{max} - x_{min}) = (19 - 13) / (71 - 13) = 6 / 58 \approx .1034$$

Applying this process across the entire feature column yields the normalized data, where the lowest value is mapped to 0 and the highest to 1:

Data Values	Normalized
13	0
16	0.0517
19	0.1034
22	0.1552
23	0.1724
38	0.4310
47	0.5862
56	0.7414
58	0.7759
63	0.8621
65	0.8966
70	0.9828
71	1

Choosing the Right Technique: Applications and Context

The decision to use **normalization** or **standardization** is dictated by the characteristics of your data and the requirements of the statistical model or algorithm you intend to use.

We typically choose **normalization** when we are performing an analysis involving multiple variables measured on fundamentally different scales, and it is crucial that all variables share the exact same range (e.g., 0 to 1). This ensures that no single feature dominates the calculation due to its inherent scale. For instance, if one feature is measured in kilograms and another in milligrams, normalization ensures they contribute equally to the distance metrics used by the algorithm. Normalization is also preferred in neural networks where bounded inputs (0 to 1) are often necessary for activation functions.

Conversely, we generally opt for **standardization** when the goal is to understand how individual data points relate to the central tendency and variance of the distribution. Standardization is vital for algorithms that assume features are normally distributed, such as Linear Regression, Logistic Regression, and some clustering methods. Furthermore, because standardization is based on the mean and standard deviation, it is less impacted by extreme outliers than Min-Max normalization.

For example, if you have a list of exam scores and want to determine how much better (or worse) a student performed relative to the average, standardizing the raw scores allows you to calculate the precise number of standard deviations the score lies from the mean. A standardized score of 1.26 immediately tells you that the student's score is 1.26 standard deviations above the average

performance, providing context that raw data cannot.

Key Differences Summarized

When deciding which scaling method to apply, it is helpful to keep these fundamental outcomes in mind:

A **normalized dataset** (Min-Max scaling) will always have values bounded between 0 and 1 (or any defined minimum and maximum), resulting in a fixed range.

A **standardized dataset** (Z-score scaling) will always have a mean of 0 and a standard deviation of 1, but its minimum and maximum values are not fixed and can extend far beyond the 0 and 1 bounds if outliers are present.

Ultimately, the choice between normalization and standardization depends entirely on the downstream application, the structure of your data, and how sensitive your chosen algorithm is to the scale and distribution of input features.

Related Tutorials

The following tutorials explain how to standardize and normalize data in different statistical software packages and programming environments: