

# Which Metric Should You Use: MSE or RMSE?

Authored by  
**stats writer**

May 12, 2024

## RECOMMENDED CITATION

stats writer (2024). *Which Metric Should You Use: MSE or RMSE?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=143953>

When evaluating the performance of a regression model, it is important to choose the appropriate metric to accurately measure the model's predictive ability. Two commonly used metrics are Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). MSE measures the average squared difference between the actual and predicted values, while RMSE takes the square root of MSE to provide a more interpretable measure. The decision of which metric to use depends on the specific application and the desired level of interpretability. Generally, MSE is more sensitive to outliers and larger errors, while RMSE is more intuitive and easier to interpret. Therefore, the choice between MSE and RMSE should be based on the specific needs and goals of the analysis.

## **MSE vs. RMSE: Which Metric Should You Use?**

**Regression models are used to quantify the relationship between one or more predictor variables and a .**

**Whenever we fit a regression model, we want to understand how well the model is able to use the values of the predictor variables to predict the value of the response variable.**

**Two metrics we often use to quantify how well a model fits a dataset are the mean squared error (MSE) and the root mean squared error (RMSE), which are calculated as follows:**

**MSE: A metric that tells us the average squared difference between the predicted values and the actual values in a dataset. The lower the MSE, the better a model fits a dataset.**

$$\text{MSE} = \frac{\sum(\hat{y}_i - y_i)^2}{n}$$

where:

$\Sigma$  is a symbol that means "sum"  
 $\hat{y}_i$  is the predicted value for the  $i$ th observation  
 $y_i$  is the observed value for the  $i$ th observation  
 $n$  is the sample size

**RMSE:** A metric that tells us the square root of the average squared difference between the predicted values and the actual values in a dataset. The lower the RMSE, the better a model fits a dataset.

It is calculated as:

$$\text{RMSE} = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{n}}$$

where:

$\Sigma$  is a symbol that means "sum"  
 $\hat{y}_i$  is the predicted value for the  $i$ th observation  
 $y_i$  is the observed value for the  $i$ th observation  
 $n$  is the sample size

Notice that the formulas are nearly identical. In fact, the root mean squared error is just the square root of the mean squared error.

## RMSE vs. MSE: Which Metric Should You Use?

When assessing how well a model fits a dataset, we use the RMSE more often because it is measured in the same units as the response variable.

Conversely, the MSE is measured in squared units of the response variable.

To illustrate this, suppose we use a regression model to predict the number of points that 10 players will score in a basketball game.

| Predicted Points ( $\hat{y}_i$ ) | Actual points ( $y_i$ ) |
|----------------------------------|-------------------------|
| 14                               | 12                      |
| 15                               | 15                      |
| 18                               | 20                      |
| 19                               | 16                      |
| 25                               | 20                      |
| 18                               | 19                      |
| 12                               | 16                      |
| 12                               | 20                      |
| 15                               | 16                      |
| 22                               | 16                      |

We would calculate the mean squared error (MSE) as:

$$\text{MSE} = \frac{\sum (\hat{y}_i - y_i)^2}{n} \quad \text{MSE} =$$

$$((14-12)^2 + (15-15)^2 + (18-20)^2 + (19-16)^2 + (25-20)^2 + (18-19)^2 + (12-16)^2 + (12-20)^2 + (15-16)^2 + (22-16)^2) / 10 \text{MSE} = 16$$

The mean squared error is 16. This tells us that the average squared difference between the predicted values made by the model and the actual values is 16.

The root mean squared error (RMSE) would simply be the square root of the MSE:

$$\text{RMSE} = \sqrt{\text{MSE}} \text{RMSE} = \sqrt{16} \text{RMSE} = 4$$

The root mean squared error is 4. This tells us that the average deviation between the predicted points scored and the actual points scored is 4.

Notice that the interpretation of the root mean squared error is much more straightforward than the mean squared error because we're talking about 'points scored' as opposed to 'squared points scored.'

#### How to Use RMSE in Practice

In practice, we typically fit several regression models to a dataset and calculate the root mean squared error (RMSE) of each model.

**We then select the model with the lowest RMSE value as the "best" model because it is the one that makes predictions that are closest to the actual values from the dataset.**

**Note that we can also compare the MSE values of each model, but RMSE is more straightforward to interpret so it's used more often.**

**Additional Resources**