How to Calculate Root Mean Square Error (RMSE) to Evaluate Your Regression Model

Authored by
stats writer

December 5, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Calculate Root Mean Square Error (RMSE) to Evaluate Your Regression Model*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=105635

Regression analysis is a powerful statistical technique utilized to understand and quantify the relationship between one or more predictor variables and a response variable. When developing these models, ensuring they accurately capture the underlying data patterns is paramount. Evaluating the quality of fit requires robust metrics that quantify the typical magnitude of error associated with the predictions.

One of the most essential and widely used metrics for assessing the performance of a regression model is the **Root Mean Square Error** (RMSE). This metric provides a standardized way to measure the average magnitude of the errors, or deviations, between the predicted values generated by the model and the actual observed values in the dataset.

The lower the RMSE, the better a given model is able to "fit" a dataset. The RMSE effectively summarizes the prediction error in the same units as the response variable, making it intuitively interpretable. This metric is indispensable in statistical modeling and machine learning for comparing and selecting optimal models, as it quantifies accuracy precisely.

The Mathematical Foundation of RMSE

To fully appreciate the utility of the Root Mean Square Error, it is crucial to understand its mathematical derivation. The term **RMSE** is often used interchangeably with Root Mean Square Deviation (RMSD). It is fundamentally derived by calculating the square root of the mean of the squared residuals, which are the differences between the predicted and observed values.

The formal definition of the Root Mean Square Error, often abbreviated **RMSE**, is given by the following formula:

RMSE = $\sqrt{\Sigma(P_i - O_i)^2 / n}$

This formula encapsulates a process that systematically measures model inaccuracy. The initial squaring operation ensures that positive and negative errors do not cancel out, and also heavily penalizes larger errors, thereby signaling the presence of significant outliers or poor predictions. The components of the calculation are defined as follows:

Σ: This is the summation symbol, indicating the need to sum the subsequent squared differences across all observations.

P_i: This represents the **predicted value** generated by the model for the ith observation in the dataset.

O_i: This represents the **observed value** (the actual ground truth) for the ith observation in the dataset.

n: This is the total sample size, or the count of data points included in the calculation.

Application Example: Calculating RMSE in Practice

Let us demonstrate the practical calculation and interpretation of RMSE using a simple linear regression model. Suppose our objective is to predict the final "Exam Score" of students based on their "Hours Studied" for a college entrance examination.

We begin by collecting raw data for 15 students, recording the actual exam score and the corresponding hours studied. This data provides the necessary inputs for training the model:

Hours Studied	Exam Score
1	68
1	78
1	75
2	83
2	80
2	78
2	89
2	93
3	90
3	91
4	94
5	88
5	84
5	90
6	94

Next, using statistical software (such as R, Python, or SPSS), we fit a linear regression line to these 15 data points. This process yields a fitted regression equation that best describes the relationship between hours studied and exam score:

Exam Score = 75.95 + 3.08*(Hours Studied)

We then apply this equation to each student's hours studied to generate a predicted exam score (Pi). We can then visually compare these predicted values against the actual observed scores (Oi) to gauge the initial error:

Hours Studied	Exam Score	Predicted Score
1	68	79.03
1	78	79.03
1	75	79.03
2	83	82.11
2	80	82.11
2	78	82.11
2	89	82.11
2	93	82.11
3	90	85.19
3	91	85.19
4	94	88.27
5	88	91.35
5	84	91.35
5	90	91.35
6	94	94.43

The final step involves calculating the squared difference between the predicted and actual scores, summing these differences, dividing by the sample size (n=15), and extracting the square root, as shown in the full calculation below:

Hours Studied	Exam Score	Predicted Score	Squared Difference
1	68	79.03	121.661
1	78	79.03	1.061
1	75	79.03	16.241
2	83	82.11	0.792
2	80	82.11	4.452
2	78	82.11	16.892
2	89	82.11	47.472
2	93	82.11	118.592
3	90	85.19	23.136
3	91	85.19	33.756
4	94	88.27	32.833
5	88	91.35	11.223
5	84	91.35	54.023
5	90	91.35	1.823
6	94	94.43	0.185
		RMSE	**5.681**

The calculated RMSE for this specific regression model turns out to be **5.681**. This value indicates that, on average, our model's predictions deviate from the true exam scores by about 5.681 points.

Connecting RMSE to Residuals and Variance

The RMSE is fundamentally based on the concept of residuals. Recall that the residuals of a regression model are the differences between the observed data values and the predicted values generated by the model. These residuals represent the unexplained variance or noise in the data.

The residual calculation is straightforward:

Residual = (Pi - Oi)

where Pi is the predicted value and Oi is the observed value. By definition, the RMSE calculation squares these differences, finds the mean, and then takes the square root. The process of finding the mean of the squared differences is known as the Mean Squared Error (MSE), which is an estimate of the error variance.

This relationship means that **the RMSE represents the square root of the variance of the residuals.** In practical terms, the RMSE is the standard deviation of the prediction errors. This is a highly useful characteristic because it returns the error metric back into the original units of measurement, allowing for direct, meaningful interpretation in the context of the data being analyzed.

Using RMSE for Model Selection and Comparison

The primary strength of the RMSE lies in its utility for objective model comparison. It allows data scientists to evaluate and rank different regression models based on their actual predictive performance on the same dataset. The model with the lowest RMSE is consistently preferred, as it minimizes the average deviation from the observed data.

For instance, suppose we are testing three distinct models designed to predict a certain outcome variable. These models might use different sets of predictor variables or employ different types of regression algorithms (e.g., Ridge vs. Lasso). We fit these models and calculate their respective RMSE values:

RMSE of Model 1: **14.5**

RMSE of Model 2: **16.7**

RMSE of Model 3: **9.8**

Model 3, with an RMSE of 9.8, clearly demonstrates the best fit among the three candidates, as its predictions exhibit the smallest average error. This immediate quantitative comparison is why RMSE is favored over metrics that are less sensitive to absolute error magnitude.

RMSE vs. R-Squared: Understanding the Differences

When evaluating model performance, RMSE is often contrasted with the coefficient of determination, or R-squared (R^2). Although both assess model quality, they serve fundamentally different purposes and convey unique information.

As discussed, the RMSE is an absolute measure of error, expressed in the same units as the response variable. It focuses strictly on quantifying the prediction accuracy and is highly sensitive to large prediction errors due to the squaring of the residuals.

In contrast, R-squared is a relative measure. It reports the proportion of the total variance in the dependent variable that is successfully explained by the predictor variables in the model. R-squared values range from 0 to 1; a value close to 1 indicates that the model explains a high percentage of the variability.

The key distinction is interpretability: RMSE tells us "how wrong" the predictions are on average, in real-world terms (e.g., 5.681 points off). R-squared tells us "how much variability" the model accounts for (e.g., 85% of variance explained). Both are essential for a complete assessment of any robust predictive model.