Root-Mean-Square-Deviation/Error (RMSD/RMSE)

The **root-mean-square-deviation (RMSD)** or **root-mean-square-error (RMSE)** is a frequently used measure of the differences between values (sample or population) predicted by a model or an **estimator** and the values observed. It represents the square root of the second **sample moment** of the differences between predicted values and observed values or the **quadratic mean** of these differences. These **deviations** are called **residuals** when the calculations are preformed over the data sample that was used for estimation and are called *errors* (or prediction errors) when computed out-of-sample.

Data:

actual: (x, y)	predicted: (x, y)	$\bar{\mathbf{x}} = 2$	$\bar{\mathbf{y}} = 3$
$(1, 1)$	$(1, 0.5)$	$\mathbf{s_x} = 0.816$	$\mathbf{s_y} = 2.160$
$(2, 2)$	$(2, 3)$		
$(2, 3)$	$(2, 3)$		
$(3, 6)$	$(3, 5.5)$	(Prediction Formula: $\hat{y} = 2.50x - 2$)	

Insight: The $\mathbf{i^{th}}$ residual will be equal to the $\mathbf{i^{th}}$ y-value for a given x, minus the predicted y-value for a given x:

$$r_i = y_i - \hat{y} \tag{1}$$

Calculate Residuals:

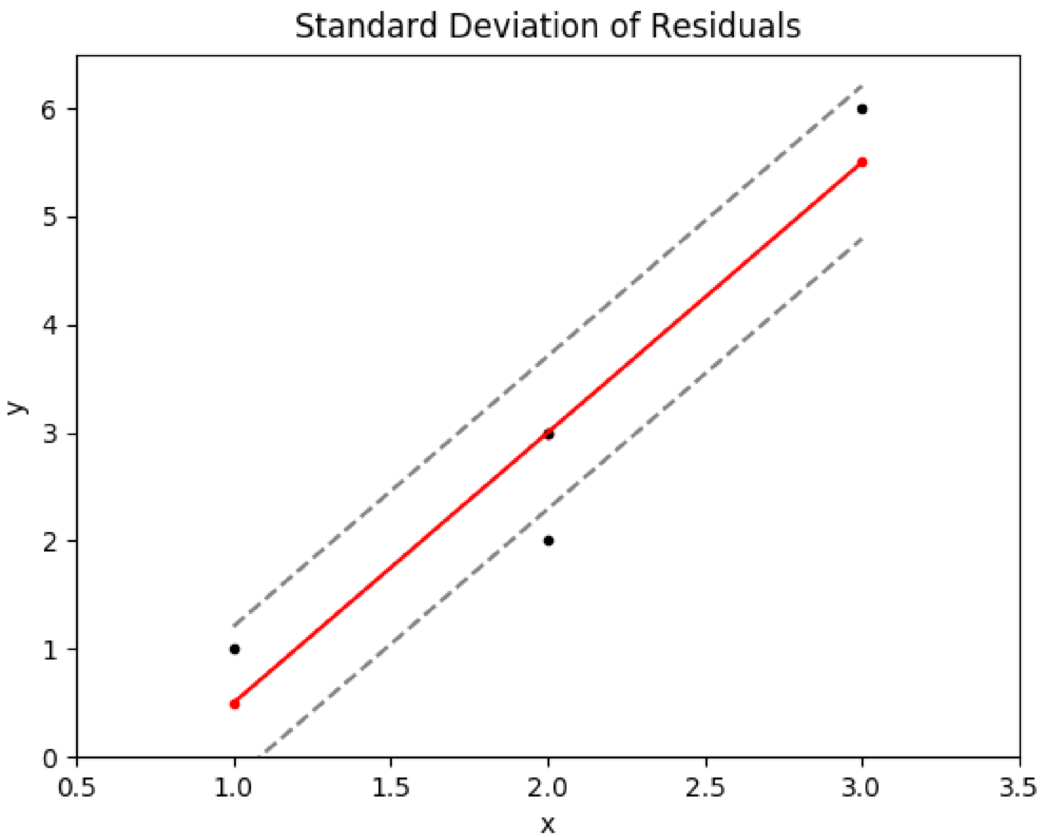
$$\begin{aligned} r_1 &= 1 - 0.5 & &\Rightarrow & r_1 &= 0.5 \\ r_2 &= 2 - 3 & &\Rightarrow & r_2 &= -1 \\ r_3 &= 3 - 3 & &\Rightarrow & r_3 &= 0 \\ r_4 &= 6 - 5.5 & &\Rightarrow & r_4 &= 0.5 \end{aligned}$$

Insight: Similar to typical standard deviation, but take the distance between a point and the model's **prediction**, sum the result, and like a sample standard deviation, divide by $n - 1$, then take the square-root of the result.

$$\sqrt{\frac{\sum r_i}{n-1}} \tag{2}$$

Calculate Standard Deviation of Residuals:

$$\sqrt{\frac{(0.5)^2 + (-1)^2 + (0)^2 + (0.5)^2}{3}} = \sqrt{\frac{(0.25) + (1) + (0) + (0.25)}{3}} = \sqrt{\frac{1.5}{3}} = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}} \approx \boxed{0.707}$$



Summary: This is used to find out how much a model disagrees with the actual data. A **lower number** means a **better** fit to the model.