What are standardized residuals?

Authored by

stats writer

December 15, 2025

ARABPSYCHOLOGY.COM

RECOMMENDED CITATION

stats writer (2025). *What are standardized residuals?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=107487

Standardized residuals are a crucial tool in statistical modeling, particularly in assessing the assumptions and identifying influential observations within linear regression analysis. Unlike raw residuals, which measure the simple difference between observed and predicted values, standardized residuals adjust this difference by accounting for the inherent variability of the observation. This transformation ensures that all residuals are placed onto a common scale, making them directly comparable regardless of their position relative to the mean of the predictor variables. By normalizing these values, we can confidently apply universal thresholds--such as an absolute value of 2 or 3--to flag potential outliers or points that may be unduly influencing the fitted model.

The core benefit of standardization lies in mitigating the effect of varying precision across different data points. In regression analysis, the variance of residuals is generally not constant; residuals associated with observations having high leverage (i.e., extreme predictor values) tend to have smaller variances. If we only examine raw residuals, an observation with a potentially meaningful deviation might appear less severe simply because it has higher leverage, thereby masking a genuine outlier. Standardized residuals correct for this issue by dividing the raw residual by its estimated standard deviation, thus ensuring that the magnitude of the resulting value accurately reflects the point's deviance relative to what is expected at that specific predictor level.

Understanding the Foundation: What is a Residual?

Before delving into the technicalities of standardization, it is essential to firmly grasp the concept of a residual. In statistical modeling, particularly regression analysis, a residual represents the error associated with a single data point. It is the fundamental measure of how well the calculated regression line predicts the observed outcome for that specific case. Mathematically, the residual (e_i) is the difference between the observed value (Y_i) and the predicted value ($\hat{Y}_i$):

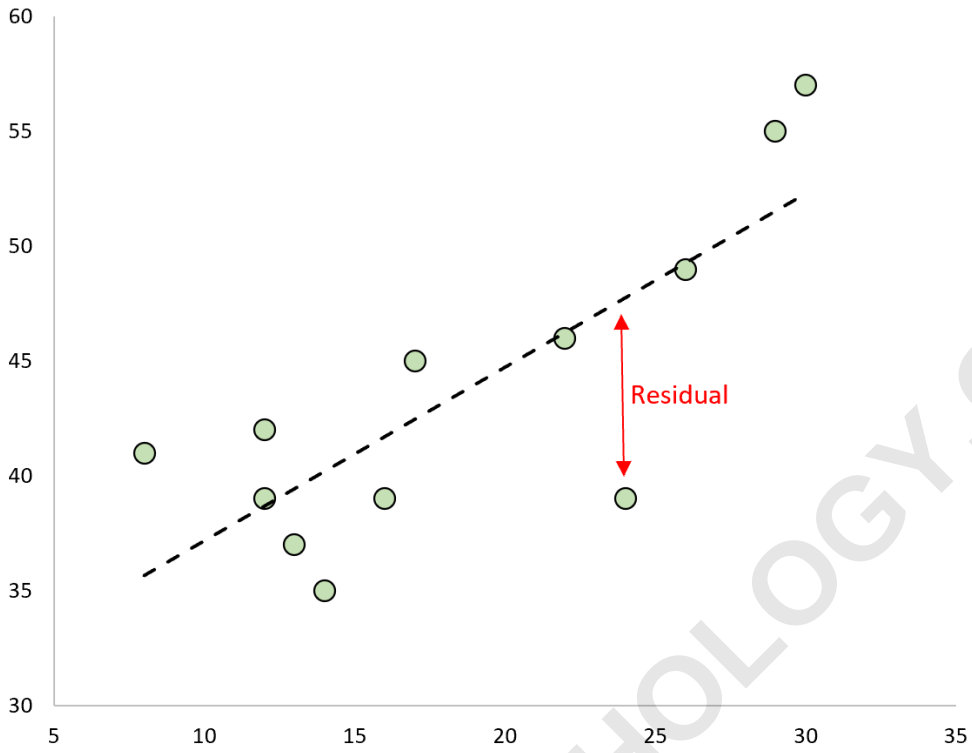
Residual = Observed Value - Predicted Value

When visualizing a regression model on a scatterplot, the residual has a clear geometric interpretation. For any given observation, the residual is the vertical distance separating the actual data point from the fitted regression line. A positive residual indicates that the model underestimates the observed value, while a negative residual means the model overestimates it. Residuals close to zero suggest that the model provides an excellent fit for that particular observation.

The quality and distribution of these residuals are paramount to validating the overall regression model. Assumptions of linear regression, such as the normality of errors and homoscedasticity (constant variance of errors), are checked primarily through analyzing the residuals. However, using the raw residual values alone can be misleading when attempting to detect true outliers, which is where the standardization process becomes indispensable.

ARABPSYCHOLOGY.COM

If we plot the observed values and overlay the fitted regression line, the residuals for each observation would be the vertical distance between the observation and the regression line:



The Problem with Raw Residuals: Non-Constant Variance

While raw residuals provide a straightforward measure of prediction error, their variance is not constant across all observations, which violates one of the core principles often desired in statistical diagnostics. The precision of the residual estimate depends heavily on the leverage of the observation. Leverage (h_{ii}) quantifies how far an observation's predictor value is from the mean of all predictor values. Observations with high leverage exert greater influence on the slope of the regression line and, consequently, their residuals tend to be smaller and less variable.

Because high-leverage points essentially "pull" the line closer to themselves, their raw residuals will naturally possess a smaller variance compared to residuals from observations located nearer to the average of the predictor values. This variation in residual variance means that a raw residual of, say, 10 units might represent a significant deviation if it comes from a high-leverage point, but only a modest deviation if it comes from a low-leverage point. This inconsistency makes it impossible to define a single, fixed threshold for raw residuals that reliably identifies unusual points across the entire dataset.

To overcome this limitation, we must scale the raw residual by its individual standard deviation. This scaling process is precisely what yields the standardized residual. By adjusting for the unique

ARABPSYCHOLOGY.COM

variability associated with each observation's position, we effectively create a homogeneous set of errors that can be compared directly, allowing us to assess unusualness in a standardized, meaningful way, much like a z-score standardizes a raw score relative to the mean and standard deviation of a sample.

The Formal Calculation of the Standardized Residual

The standardized residual, often denoted as r_i, is calculated by taking the raw residual (e_i) and dividing it by an estimate of its standard deviation. This adjustment factor is critical because it incorporates both the model's overall error and the observation's specific leverage. The standard deviation of the i^{th} residual, $s(e_i)$, is calculated using the model's residual standard error (RSE) and the observation's leverage (h_{ii}).

The complete formula for the standardized residual is:

ri = ei / s(ei)

Which is typically written out using the estimated components:

ri = ei / RSE√1-hii

Where the components are defined as follows:

ei: The raw residual for the i^{th} observation.
RSE: The Residual Standard Error of the model, which is an estimate of the standard deviation of the error term ($sigma$).
hii: The leverage of the i^{th} observation, quantifying its influence on the fitted line.

It is important to note that standardized residuals are sometimes referred to by the more technical term, "internally studentized residuals." They are distinguished from "externally studentized residuals" (or just studentized residuals), which use a residual standard error calculated from a model where the i^{th} observation has been entirely excluded. While both are used for outlier detection, the internally studentized (standardized) residual is slightly less sensitive to the observation it is calculating, whereas the externally studentized residual is designed to maximize the detection of that specific point's influence.

Interpreting Standardized Residuals and Identifying Outliers

The primary application of standardized residuals is the robust identification of unusual observations, or outliers. Because standardized residuals are scaled to account for varying variances, their distribution closely approximates a standard normal distribution, especially for large sample sizes. This allows statisticians to use probabilistic thresholds for flagging points that

ARABPSYCHOLOGY.COM

are unusually far from the regression line.

In practice, common rules of thumb for identifying potential outliers based on standardized residuals are:

Any observation with a standardized residual having an absolute value greater than 3 is generally considered a severe outlier.
Some researchers or fields may use a more conservative threshold, considering any observation with an absolute value greater than 2 to be worth investigating.

When an observation generates a standardized residual exceeding these thresholds, it serves as a strong signal that the observation is poorly predicted by the model relative to its expected variability. However, finding such a point does not automatically mean it should be removed. Instead, it necessitates a careful investigation. Analysts must determine if the unusual residual is the result of a measurement error, a data entry mistake, or if it genuinely represents a phenomenon not captured by the current model structure. Ignoring or automatically removing outliers without investigation can lead to biased results and loss of important information.

Example: Step-by-Step Calculation of Standardized Residuals

To illustrate the process, let us consider a small dataset consisting of 12 observations, with a single predictor X and an outcome Y.

Suppose we have the following dataset with 12 total observations:

X	Y
8	41
12	42
12	39
13	37
14	35
16	39
17	45
22	46
24	39
26	49
29	55
30	57

Using standard statistical software (such as R, Excel, Python, or Stata), we first fit a linear regression model to this data. For this example, assume the line of best fit derived from the

ARABPSYCHOLOGY.COM

software is:

$Y = 29.63 + 0.7553X$

Step 1: Calculate Predicted Values and Raw Residuals

We use the fitted equation to calculate the predicted Y value ($hat{Y}_i$) for each observation based on its X value. For instance, for the first observation, where $X=8$:

$hat{Y}_1 = 29.63 + 0.7553 times (8) = mathbf{35.67}$

The raw residual (e_1) is then found by subtracting this predicted value from the observed value ($Y_1 = 41$):

Residual = Observed value - Predicted value = $41 - 35.67 = mathbf{5.33}$

We repeat this calculation for all 12 observations to determine the full set of raw residuals:

X	Y	Predicted Value	Residual
8	41	35.67	5.33
12	42	38.69	3.31
12	39	38.69	0.31
13	37	39.45	-2.45
14	35	40.20	-5.20
16	39	41.71	-2.71
17	45	42.47	2.53
22	46	46.25	-0.25
24	39	47.76	-8.76
26	49	49.27	-0.27
29	55	51.53	3.47
30	57	52.29	4.71

ARABPSYCHOLOGY.COM

Step 2: Determine RSE and Leverage (h_{ii})

To standardize these raw residuals, we require two additional pieces of information obtained from the statistical software: the Residual Standard Error (RSE) and the leverage (h_{ii}) for each point. For this specific model, let us assume the software calculates the RSE as $mathbf{4.44}$.

The leverage statistic (h_{ii}) is a matrix-based calculation that is usually handled entirely by statistical software. It quantifies the influence of each X-value on the fitted line. Although the mathematical derivation of leverage is complex, we use the values provided by the software:

X	Y	Predicted Value	Residual	h_{ii}
8	41	35.67	5.33	0.27
12	42	38.69	3.31	0.16
12	39	38.69	0.31	0.16
13	37	39.45	-2.45	0.14
14	35	40.20	-5.20	0.12
16	39	41.71	-2.71	0.09
17	45	42.47	2.53	0.09
22	46	46.25	-0.25	0.10
24	39	47.76	-8.76	0.13
26	49	49.27	-0.27	0.18
29	55	51.53	3.47	0.27
30	57	52.29	4.71	0.30

Step 3: Calculate the Standardized Residual (r_i)

With the raw residual (e_i), the RSE, and the leverage (h_{ii}) determined, we can now apply the standardization formula:

ri = ei / RSE√1-hii

For the first observation, $e_1 = 5.33$, $RSE = 4.44$, and $h_{11} = 0.27$. The standardized residual is calculated as:

r1 = 5.33 / (4.44 $times$ $sqrt{1 - 0.27}$) = 5.33 / (4.44 $times$ $sqrt{0.73}$) $approx$ 5.33 / 3.799 $approx$ 1.404

Repeating this rigorous process for every observation yields the final set of standardized residuals:

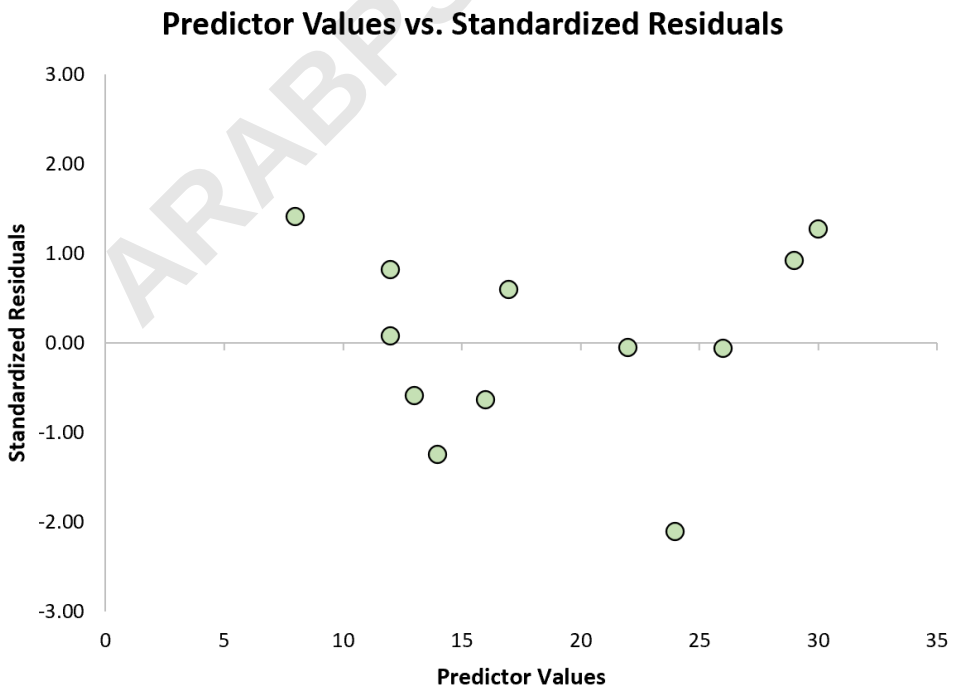
ARABPSYCHOLOGY.COM

X	Y	Predicted Value	Residual	h_{ii}	Standardized Residual
8	41	35.67	5.33	0.27	1.40
12	42	38.69	3.31	0.16	0.81
12	39	38.69	0.31	0.16	0.08
13	37	39.45	-2.45	0.14	-0.59
14	35	40.20	-5.20	0.12	-1.25
16	39	41.71	-2.71	0.09	-0.64
17	45	42.47	2.53	0.09	0.60
22	46	46.25	-0.25	0.10	-0.06
24	39	47.76	-8.76	0.13	-2.11
26	49	49.27	-0.27	0.18	-0.07
29	55	51.53	3.47	0.27	0.91
30	57	52.29	4.71	0.30	1.27

Visualizing and Concluding the Analysis

The final and perhaps most practical step in utilizing standardized residuals is plotting them against the predictor values (or sometimes the fitted values) to visually assess the model diagnostics. This graphical representation allows for quick identification of points that breach the standard thresholds and helps verify the assumption of homoscedasticity.

We can create a scatterplot of the predictor values (X) versus the calculated standardized residuals (r_i) and overlay reference lines at +3 and -3 (and sometimes +2 and -2):



ARABPSYCHOLOGY.COM

Upon reviewing the plot, we observe that none of the standardized residuals exceed an absolute value threshold of 3. In fact, the largest standardized residual is 1.404, which is well within the acceptable range. Based on this analysis, we conclude that there are no severe outliers present in the dataset that unduly affect the regression model fit, suggesting the model is a reasonable fit for the majority of the data points.

While the threshold of 3 is often the universally accepted standard for outlier identification, some researchers, particularly in fields requiring tighter controls, may opt to investigate points exceeding an absolute value of 2. The choice between using a threshold of 2 or 3 is often context-dependent, relying on industry standards, sample size, and the perceived consequences of including or excluding an influential observation. Regardless of the chosen threshold, standardized residuals provide the necessary statistical metric to conduct this critical diagnostic assessment reliably.

ARABPSYCHOLOGY.COM