

Understanding Null and Residual Deviance in Generalized Linear Models

Authored by
Mohammed Iooti

November 2, 2025

RECOMMENDED CITATION

Mohammed Iooti (2025). *Understanding Null and Residual Deviance in Generalized Linear Models*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=8774>

When constructing statistical models, particularly those falling under the umbrella of a [Generalized Linear Model \(GLM\)](#)--such as [logistic regression](#) or [Poisson regression](#)--analysts must assess how well the chosen model describes the observed data. Statistical software provides two essential metrics for this assessment: the **null deviance** and the **residual deviance**. These values are paramount for determining the model's overall utility and its goodness-of-fit.

In the framework of GLMs, deviance serves as a measure of discrepancy, analogous to the sum of squared errors used in traditional linear regression. It quantifies the lack of fit by comparing the likelihood of the model being tested against a hypothetical, perfect model known as the [saturated model](#). Interpreting the difference between the null and residual deviances is the definitive method for assessing whether your chosen set of predictor variables contributes meaningful explanatory power to the analysis.

Defining Deviance: The Goodness-of-Fit Metric

Before diving into the specifics of null and residual values, it is crucial to solidify the concept of deviance itself. Deviance is formally defined as twice the difference between the log-likelihood of the maximum achievable model (the saturated model) and the log-likelihood of the current model being evaluated. Since the saturated model always perfectly fits the data and thus has a deviance of zero, the calculated deviance for any other model provides a scale for measuring how far that model is from perfection.

GLMs often involve non-normal error distributions (like binomial or Poisson), making traditional measures like R-squared less applicable or harder to interpret. Deviance provides a standardized, likelihood-based method for evaluating model performance across various distribution families. A model with excellent fit will have a deviance value close to zero, signifying a strong alignment between the predicted outcomes and the actual observed outcomes.

Understanding Null Deviance: Establishing the Baseline

The [Null Deviance](#) is the foundational benchmark against which all more complex models must be compared. It represents the deviance of the simplest possible model that can be fitted to the data. This baseline model includes only an [intercept term](#) and excludes all other predictor variables (covariates).

The intercept-only model essentially postulates that all observed outcomes are best predicted by the overall mean of the response variable (or, more precisely, the mean transformed via the link function). Therefore, the Null Deviance indicates the total inherent variability within the response variable that must be explained by any meaningful predictive model. It sets the maximum possible reduction in deviance achievable by incorporating the full set of available predictors.

The degrees of freedom associated with the Null Deviance are calculated as $N - 1$, where N represents the total number of observations in the dataset. This accounts for the single parameter (the intercept) that has been estimated in this minimal model. This metric serves as the starting point for calculating overall model improvement.

Interpreting Residual Deviance and Measuring Fit

The [Residual Deviance](#) provides the ultimate measure of the lack of fit remaining after the full, proposed model--incorporating all p predictor variables--has been estimated. It compares the likelihood of your fully specified model against the likelihood of the saturated model. In essence, it tells you how much unexplained variability is left over after accounting for the variables you have included.

The primary goal of statistical model fitting is to achieve a substantial reduction in deviance. Consequently, a **lower Residual Deviance value** signifies superior model performance, indicating that the chosen predictors have successfully accounted for a large proportion of the total variability initially captured by the Null Deviance. If the Residual Deviance is very large, it suggests that the model is poorly specified or that important variables have been omitted.

The degrees of freedom (df) for the Residual Deviance are calculated as $N - p - 1$, where N is the number of observations and p is the number of predictor variables utilized. The drop in deviance from the Null Model to the Residual Model is the core measure of the gain in explanatory power achieved by introducing the covariates.

Testing Overall Model Utility: The Likelihood Ratio Test (LRT)

While observing a reduction in deviance is informative, statistical rigor demands a formal test to confirm that this improvement is significant and not merely due to chance. This is achieved using the [Likelihood Ratio Test \(LRT\)](#), which is specifically designed to assess the overall significance of the predictors collectively.

The LRT compares the intercept-only model (the null hypothesis) against the full model (the alternative hypothesis). The test statistic, often denoted as X^2 or G^2 , is derived simply by finding the difference between the two primary deviance metrics:

$$X^2 = \text{Null deviance} - \text{Residual deviance}$$

This calculated test statistic approximates a [Chi-Square statistic](#). This X^2 value is then evaluated against a Chi-Square distribution where the degrees of freedom (df) are equal to the difference between the Null df and the Residual df. Crucially, this difference always equates to p , the exact number of predictor variables added to the model. The larger the resulting X^2 value,

the greater the evidence of significant model improvement.

The final step involves determining the [p-value](#) associated with this Chi-Square statistic. If this p-value falls below the predefined significance level (commonly $\alpha = 0.05$), we confidently reject the null hypothesis. This rejection allows us to conclude that the full model, incorporating the predictors, provides a statistically superior and significantly better fit to the data than the basic intercept-only model.

Practical Example: Assessing Model Fit using Logistic Regression in R

To fully grasp the interpretation of these concepts, we will walk through a practical demonstration using a [logistic regression](#) model implemented in the R statistical environment. Our example utilizes the widely recognized **Default** dataset, sourced from the [ISLR package](#), which contains attributes pertinent to individuals' credit default status.

We begin by loading the necessary data and executing a preliminary summary inspection to familiarize ourselves with the variables and their distributions:

```
#load dataset
```

```
data <- ISLR::Default
```

```
#view summary of dataset
```

```
summary(data)
```

```
default student balance income
```

```
No :9667 No :7056 Min. : 0.0 Min. : 772
```

```
Yes: 333 Yes:2944 1st Qu.: 481.7 1st Qu.:21340
```

```
Median : 823.6 Median :34553
```

```
Mean : 835.4 Mean :33517
```

```
3rd Qu.:1166.3 3rd Qu.:43808
```

```
Max. :2654.3 Max. :73554
```

This dataset tracks 10,000 observations and includes the following attributes relevant to predicting credit default:

default: The binary response variable (Yes/No), indicating whether an individual defaulted.

student: A categorical factor representing student status.

balance: A continuous measure of the individual's average bank balance.

income: A continuous measure detailing the individual's annual income.

Our objective is to fit a logistic regression model using `balance`, `student`, and `income` as

predictors to estimate the probability of `default`. The resulting model summary output immediately provides the critical deviance statistics we require for the Likelihood Ratio Test:

#fit logistic regression model

```
model <- glm(default~balance+student+income, family="binomial", data=data)
```

```
#view model summary
summary(model)
```

Call:

```
glm(formula = default ~ balance + student + income, family = "binomial",
data = data)
```

Deviance Residuals:

Min 1Q Median 3Q Max

-2.4691 -0.1418 -0.0557 -0.0203 3.7383

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -1.087e+01 4.923e-01 -22.080 < 2e-16 ***

balance 5.737e-03 2.319e-04 24.738 < 2e-16 ***

studentYes -6.468e-01 2.363e-01 -2.738 0.00619 **

income 3.033e-06 8.203e-06 0.370 0.71152

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2920.6 on 9999 degrees of freedom

Residual deviance: 1571.5 on 9996 degrees of freedom

AIC: 1579.5

Number of Fisher Scoring iterations: 8

Interpreting the Results and Conclusion

From the R summary output, we extract the two critical figures required for our assessment of overall model fitness:

Null deviance: 2920.6 with df = 9999

Residual deviance: 1571.5 with df = 9996

We observe that three predictor variables (`balance`, `student`, and `income`) were included in the full model, meaning we have $p = 3$ degrees of freedom for our comparative test. This is mathematically confirmed by the difference in the degrees of freedom reported: $9999 - 9996 = 3$.

Next, we calculate the Chi-Square test statistic (X^2) by quantifying the substantial reduction in deviance achieved by moving from the intercept-only model to the full model:

$X^2 = \text{Null deviance} - \text{Residual deviance}$

$X^2 = 2920.6 - 1571.5$

$X^2 = 1349.1$

We then compare this extremely large X^2 value (1349.1) against a [Chi-Square distribution](#) with 3 degrees of freedom. Given the massive magnitude of the reduction in deviance (a reduction of approximately 1350 units), the resulting [p-value](#) is infinitesimally small--many orders of magnitude below the standard significance threshold of 0.05 .

The statistical conclusion is unambiguous: we strongly reject the null hypothesis that the predictors have no effect. We conclude that the fitted model, which incorporates bank balance, student status, and income, is **highly effective and statistically useful** for predicting an individual's probability of default. The predictors collectively explain a highly significant portion of the variability in default status that the simple baseline model could not account for.

Summary of Key Interpretations

To summarize the fundamental distinction between the two deviance metrics:

The [Null Deviance](#) represents the total variation to be explained; it is the starting point.

The [Residual Deviance](#) represents the variation remaining unexplained after fitting the full model; it is the endpoint.

The difference between them, tested via the [Likelihood Ratio Test](#), determines whether the model provides a significant improvement over the baseline.

Mastering the interpretation of null and residual deviance is a foundational skill for anyone working with [Generalized Linear Models](#), providing a robust, likelihood-based measure of model performance that transcends the constraints of standard linear regression assumptions.