

How to Easily Calculate AIC in R for Model Comparison

Authored by
stats writer

December 5, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Calculate AIC in R for Model Comparison*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=105529>

The Akaike Information Criterion (AIC) is a fundamental concept in statistical modeling, serving as an invaluable tool for comparing the relative quality of multiple candidate statistical models. Developed by statistician Hirotugu Akaike, AIC provides a standardized method for evaluating how well a model fits the data while simultaneously penalizing the model for excessive complexity. This principle is vital because a model with too many parameters might achieve a deceptively excellent fit on the training data (overfitting), but fail miserably when applied to new, unseen observations. By incorporating both the goodness-of-fit and the number of estimated parameters, AIC helps researchers identify the most parsimonious model--the one that achieves the best balance between explanatory power and simplicity. When working in the R statistical environment, the calculation of this metric is straightforward, primarily utilizing the built-in `AIC()` function for single model analysis or, more frequently for model comparison, specialized functions from dedicated packages like `AICcmodavg`.

For standard statistical models, such as linear or generalized regression models, the `AIC()` function in R automatically takes the fitted model object--such as the result of an `lm()` call--and extracts the necessary components, including the maximum log-likelihood and the degrees of freedom, to compute the criterion. While the base function is sufficient for many scenarios, advanced model comparison often requires the small sample correction (AICc), necessitating additional functions. Understanding the nuances of when to apply AIC versus AICc, and how to interpret the resulting comparison table, is essential for robust statistical inference and model selection. Throughout this guide, we will explore the theoretical underpinnings of AIC and demonstrate the precise steps required to calculate and interpret the results using practical examples within R.

The Theoretical Foundation of AIC

The Akaike Information Criterion serves as an estimator of the relative quality of statistical models for a given set of data. It is grounded in information theory, specifically related to the Kullback-Leibler information, which measures the information lost when one model is used to approximate reality. In practice, AIC estimates the relative amount of information lost by a given model, meaning the model that minimizes the information loss is the preferred model. This minimization usually translates directly to selecting the model with the smallest AIC value among the set of candidates being compared. It is critical to note that AIC is a **relative measure**; it does not provide an absolute test of model fit, but rather allows us to rank models based on their efficiency in explaining the observed data while maintaining parsimony and minimizing the risk of overfitting.

A key strength of using AIC for model selection, particularly with regression models, is its inherent mechanism to counteract the problem of overfitting. When researchers add more predictor variables to a model, the model's goodness-of-fit (often measured by the maximum log-likelihood) will almost always improve, regardless of whether those variables contribute genuinely useful

predictive information. AIC balances this improvement against the cost associated with complexity. If adding a new parameter does not sufficiently increase the log-likelihood (the fit), the penalty term associated with the increased number of parameters will dominate, resulting in a higher, and thus less desirable, AIC value. This rigorous approach ensures that model selection is driven by substantive explanatory power rather than merely maximizing statistical noise capture.

Understanding the AIC Formula

To appreciate how this balance between fit and complexity is achieved, we must examine the core mathematical definition of the AIC. The standard formula is expressed as:

$$\text{AIC} = 2K - 2\ln(L)$$

This straightforward formula comprises two main components that define the trade-off. The first term, **2K**, represents the penalty for complexity. The second term, **-2ln(L)**, is directly related to the model's goodness-of-fit.

K (The Number of Model Parameters): This value represents the total count of estimated parameters within the model, including not only the slope coefficients for all predictor variables but also the intercept and, where applicable (such as in linear regression), the estimated variance of the error term. For instance, in a simple linear regression model with one predictor variable, K is typically 3 (one intercept, one slope coefficient, and one error variance parameter). The term 2K imposes a penalty proportional to the model's complexity, ensuring that models with more predictors are inherently penalized unless the resulting improvement in fit is substantial.

ln(L) (The Log-Likelihood): This term, specifically the maximum log-likelihood of the model, measures how probable the observed data are, given the specified model structure. A higher log-likelihood value indicates a better fit of the model to the data. Since we want to minimize the overall AIC value, the term is preceded by a negative sign (-2), meaning that a better fitting model (higher L) will yield a smaller AIC value, counteracting the complexity penalty (2K). Most statistical software, including R, calculates the maximum log-likelihood automatically when a model is fitted.

The design of the AIC formula ensures that when comparing multiple candidate models, the one possessing the **minimum AIC value** is designated as the preferred model. This model is considered the most likely candidate to provide the most accurate predictions on a new dataset, having successfully explained the most variation in the current data without incorporating excessive, unnecessary complexity.

Addressing Small Samples: The AICc Correction

While the standard AIC formula performs robustly when dealing with large sample sizes, it can

exhibit a significant bias towards selecting overly complex models when the sample size (N) is small, especially relative to the number of parameters (K). To address this critical limitation, researchers often utilize the corrected AIC, known as AICc. This correction is mathematically necessary to provide accurate model ranking when the sample size is limited, preventing the selection of models that merely capitalize on chance relationships specific to the small dataset.

The AICc is essentially an adjustment to the standard AIC that incorporates a correction factor based on the ratio of the sample size to the number of parameters. The general recommendation is to use AICc whenever the ratio of the sample size (N) to the number of parameters (K) is less than approximately 40. The AICc formula introduces an additional positive term:

$$\text{AICc} = \text{AIC} +$$

The term added to the standard AIC is the correction factor. As the sample size (N) approaches infinity, this correction factor approaches zero, causing AICc to converge with the standard AIC. Conversely, for small N, the correction factor is significant, imposing a greater penalty on complex models (those with high K) and preventing the selection of models that overfit the limited data available. Due to its superior performance in constrained environments, many practitioners recommend using AICc universally, regardless of sample size, as it introduces negligible error when N is large but provides necessary correction when N is small. The `aictab()` function, which we will use shortly, automatically provides the AICc.

Prerequisites and Base AIC Calculation in R

In the R environment, calculating AIC for a single fitted model object is straightforward using the built-in `AIC()` generic function. This function automatically recognizes standard model objects (such as those created by `lm()`, `glm()`, or `aov()`) and returns the standard AIC value based on the maximum log-likelihood. For instance, if you have fitted a linear model called `lm.object`, executing `AIC(lm.object)` will yield the result. For linear models, the degrees of freedom parameter *k* (the 2*K* component in the formula) defaults to 2, representing the standard penalty. However, when the goal is comparing several models simultaneously--which is the primary purpose of AIC--relying solely on sequential calls to `AIC()` becomes cumbersome and lacks the necessary comparative statistics, such as Delta AIC and model weights.

To perform rigorous model comparison and effectively implement the AICc correction, we must turn to specialized packages. The AICcmodavg package is the definitive tool for this purpose, providing functions designed specifically for model selection based on AIC and AICc. The function `aictab()`, in particular, organizes multiple candidate models into a single, comprehensive table, ordered by their corrected AIC values (AICc), which greatly simplifies the interpretation and comparison process. Before proceeding with the practical example, ensure you have this package installed and loaded into your R session, as it provides the essential framework for high-quality

statistical model comparison.

Practical Example: Setting Up Competing Regression Models

We will now execute a practical demonstration in R, aiming to select the best predictive model for fuel efficiency (mpg) using various combinations of variables from the renowned **mtcars** dataset. This dataset is ideally suited for demonstrating model comparison techniques because it contains numerous intercorrelated vehicle performance metrics. Our objective is to fit three distinct regression models to predict mpg, each representing a different hypothesis about the most important predictor combination.

For this example, we define the following competing candidate models, using mpg as the dependent variable in each case. These models vary in complexity, allowing us to see how the AICc penalty affects the final selection:

Model 1: A complex model incorporating four key predictors: engine displacement (disp), horsepower (hp), vehicle weight (wt), and quarter-mile time (qsec).

Model 2: A simpler model focusing on physical characteristics and performance time: engine displacement (disp) and quarter-mile time (qsec).

Model 3: A model emphasizing physical attributes: engine displacement (disp) and vehicle weight (wt).

The following R code block shows the initialization and fitting process for these three candidate linear models using the `lm()` function. Note that fitting the models is the prerequisite step before AIC calculation can commence.

#fit three models using the mtcars dataset

```
model1 <- lm(mpg ~ disp + hp + wt + qsec, data = mtcars)
```

```
model2 <- lm(mpg ~ disp + qsec, data = mtcars)
```

```
model3 <- lm(mpg ~ disp + wt, data = mtcars)
```

Calculating and Comparing AICc with `aictab()`

Once the models are fitted and stored as objects (`model1`, `model2`, `model3`), the next step is to leverage the power of the AICcmodavg package to generate a comparative table. We must first load the necessary library and then organize our fitted models into a list, along with defining descriptive names that will appear in the output table. Crucially, we use the `aictab()` function, which defaults to calculating the corrected AIC (AICc), suitable for the relatively small sample size (N=32) of the `mtcars` dataset.

The `aictab()` function takes two primary arguments: `cand.set`, which is the list containing the fitted model objects, and `modnames`, which is a vector of character strings assigning names to each model in the order they appear in the list. The resulting output table automatically sorts the models based on their AICc value, from lowest to highest, making the selection of the best model immediate and transparent. This structure is far superior to manually calculating the AIC or AICc for each model object individually, as it provides standardized metrics for comparison across the entire set of candidates.

The code below demonstrates how to load the package, define the model set and names, and execute the final comparison:

library(AICcmodavg)

```
#define list of models
```

```
models <- list(model1, model2, model3)
```

```
#specify model names
```

```
mod.names <- c('disp.hp.wt.qsec', 'disp.qsec', 'disp.wt')
```

```
#calculate AIC of each model
```

```
aictab(cand.set = models, modnames = mod.names)
```

Model selection based on AICc:

	K	AICc	Delta_AICc	AICcWt	Cum.Wt	LL
disp.hp.wt.qsec	6	162.43	0.00	0.83	0.83	-73.53
disp.wt	4	165.65	3.22	0.17	1.00	-78.08
disp.qsec	4	173.32	10.89	0.00	1.00	-81.92

Interpreting the AICc Comparison Table

The output generated by the `aictab()` function provides a wealth of information necessary for sophisticated model comparison. Because the models are sorted by their AICc value, the model listed first is automatically deemed the best fit relative to the others. However, examining the other columns is essential to understand the magnitude of evidence supporting the top model versus its competitors, allowing us to assess if competing models are substantially equivalent.

Here is a detailed breakdown of the critical columns in the output:

K: This is the total number of estimated parameters in the model, including the intercept and the error variance parameter. For Model 1 (disp.hp.wt.qsec), K=6 (1 intercept + 4 slopes + 1 error variance). For Models 2 and 3, K=4. This difference in K is what drives the complexity penalty.

AICc: This is the corrected AIC value (AIC corrected for small sample sizes). This is the primary metric for ranking. The model with the smallest AICc value (162.43 for Model 1) is the best model among the candidates.

Delta_AICc: This column represents the difference between the AICc of the current model and the AICc of the top-ranked model (which always has a Delta_AICc of 0.00). Models with a Delta_AICc value less than 2 are generally considered to have substantial support and may be included in a set of plausible models. Here, Model 3 has a Delta_AICc of 3.22, suggesting it is noticeably worse than Model 1. Model 2, with 10.89, has virtually no support.

AICcWt (Akaike Weights): This is arguably the most informative metric. The AICc weight represents the probability that the given model is the best model (i.e., the one minimizing the information loss), relative to the entire set of candidate models considered. Model 1 has a weight of 0.83, meaning there is an 83% probability that this model is the most appropriate among the three tested. Model 3 has a 17% probability.

Cum.Wt: The cumulative sum of the AIC weights, calculated by summing the weights from the top model down. This is useful for identifying a confidence set of models. Here, Model 1 and Model 3 cumulatively account for 100% of the predictive power within the set.

LL: The log-likelihood of the model. This value reflects the model's goodness-of-fit to the data. Note that Model 1 has the highest (least negative) log-likelihood (-73.53), reflecting its superior fit, which was substantial enough to overcome the penalty imposed by its two extra parameters (K=6 vs K=4).

Model Selection and Final Interpretation

Based on the interpreted results, the model `disp.hp.wt.qsec` (Model 1) clearly outperforms the other candidates. It possesses the lowest AICc score (162.43) and carries an overwhelming Akaike Weight (0.83). This provides strong statistical evidence that the combination of engine displacement, horsepower, weight, and quarter-mile time provides the most efficient and parsimonious explanation for the variation in fuel economy (mpg) within this dataset. The evidence against the next best model (`disp.wt`) is substantial (Delta_AICc = 3.22).

We can translate this chosen model back into its corresponding mathematical relationship, providing a clear summary of the variables identified as most important for predicting fuel economy:

$$\text{mpg} = \beta_0 + \beta_1(\text{disp}) + \beta_2(\text{hp}) + \beta_3(\text{wt}) + \beta_4(\text{qsec}) + \varepsilon$$

Identifying the best model through AIC is only the first part of the process. Once Model 1 is selected, the researcher must proceed with a comprehensive analysis of this specific model. This

involves examining the statistical significance of the individual coefficient estimates (β values), interpreting the direction and magnitude of the relationships between the predictors and the outcome variable, and assessing overall model diagnostics, such as linearity assumptions and residual analysis. Final checks often include examining metrics like the R-squared value to understand the proportion of variance explained and ensuring the model meets standard statistical assumptions before drawing definitive conclusions about the underlying relationship between the set of predictor variables and the response.

ARABPSYCHOLOGY.COM