

# What is a Nested Model? (Definition & Example)

Authored by  
**Mohammed loot**

November 1, 2025

## RECOMMENDED CITATION

Mohammed loot (2025). *What is a Nested Model? (Definition & Example)*.  
PSYCHOLOGICAL STATISTICS. Retrieved from  
<https://statistics.arabpsychology.com/?p=7629>

## The Foundation of Nested Models in Statistical Modeling

The concept of a **nested model** is absolutely central to robust [statistical model](#) building and effective model comparison, particularly within the field of [regression analysis](#). Formally, a statistical model (Model B) is defined as nested within a larger, more comprehensive model (Model A) if the set of its independent variables--or [predictor variables](#)--is a strict subset of the variables utilized in Model A. This hierarchical structure provides the necessary mathematical framework for systematic and rigorous evaluation of model efficacy.

The defining characteristic of nesting is operational: the simpler model, often referred to as the reduced model, must be obtainable directly from the more complex model (the full model) by imposing specific constraints. These constraints typically involve setting one or more of the parameters (coefficients) associated with the extra variables in the full model precisely to zero. If this mathematical derivation is impossible--meaning the reduced model cannot be formed by nullifying coefficients in the full model--then the two models are fundamentally non-nested.

Grasping this hierarchical relationship empowers researchers and data scientists to systematically assess the incremental value of adding complexity. By comparing a full model against its nested counterpart, one can determine whether the inclusion of additional [predictor variables](#) yields a statistically meaningful improvement in explanatory power or predictive accuracy. This systematic approach is vital for achieving the goal of model parsimony--finding the simplest model that adequately explains the data.

### Practical Illustration: Building Nested Regression Models

To solidify the theoretical definition, let us explore a concrete, practical example drawn from sports analytics--specifically, predicting the performance of a basketball player. Our objective is to predict the number of points scored by a player during a game based on various performance metrics and physical attributes.

We begin by formulating the most extensive version, the **full model** (designated as Model A). This model incorporates four distinct factors hypothesized to influence scoring output. The variables included are minutes played, player height, playing position, and the number of shots attempted. Mathematically, this model is expressed as:

$$\text{Points} = \beta_0 + \beta_1(\text{minutes}) + \beta_2(\text{height}) + \beta_3(\text{position}) + \beta_4(\text{shots}) + \varepsilon$$

In this regression formula,  $\beta_0$  is the intercept term;  $\beta_1$  through  $\beta_4$  represent the estimated coefficients or weights assigned to each variable; and  $\varepsilon$  accounts for the residual error. Model A, therefore, uses all four identified [predictor variables](#) simultaneously to generate its prediction of the outcome variable (Points).

A **nested model** (Model B), also known as the reduced model, is then deliberately constructed by utilizing only a proper subset of the variables found in Model A. For example, we might hypothesize that only playing time and physical stature are truly necessary, leading to the simpler Model B:

$$\text{Points} = \beta_0 + \beta_1(\text{minutes}) + \beta_2(\text{height}) + \varepsilon$$

Crucially, **Model B is nested within Model A** because its set of predictors is entirely contained within Model A's set. This nesting is confirmed by the fundamental operational test: Model B is mathematically identical to Model A if the coefficients for the excluded variables--namely  $\beta_3$  (position) and  $\beta_4$  (shots attempted)--are constrained to zero. This constraint rule is the definitive criterion used in formal statistical comparison tests.

## The Distinction Between Nested and Non-Nested Model Structures

While understanding nesting is key, it is equally vital for practitioners to identify situations where models are fundamentally non-nested. Non-nesting occurs when the variable sets are distinct and overlapping, meaning neither model's set of predictors is a proper subset of the other. Consequently, each model incorporates unique explanatory factors that cannot be eliminated from the alternative model simply by setting coefficients to zero.

To illustrate this, let us introduce a third formulation, Model C. This model attempts to predict the same outcome (Points) but substitutes some predictors from Model A with a new, distinct variable:

$$\text{Points} = \beta_0 + \beta_1(\text{minutes}) + \beta_2(\text{height}) + \beta_3(\text{free throws attempted}) + \varepsilon$$

When we compare this new [statistical model](#) (Model C) against our original full model (Model A), the non-nested relationship becomes immediately clear. Model A includes 'position' and 'shots attempted', which are entirely absent in Model C. Conversely, Model C includes 'free throws attempted', a variable that is not present in Model A. The inclusion of unique variables in both directions violates the subset requirement.

The conclusion is that **Model C is not nested in Model A**, nor is the reverse true. This designation has profound implications for the method of comparison. Because specialized tests like the F-test or Likelihood Ratio Test rely on the specific constraints inherent in nested structures, non-nested models require alternative comparison strategies. Researchers must instead rely on techniques such as evaluating information criteria, like the [Akaike Information Criterion \(AIC\)](#) or the Bayesian Information Criterion (BIC), which are designed to compare models with non-hierarchical relationships.

## Leveraging Nested Structures for Rigorous Model Selection

The true power of nested models in applied statistics emerges through their crucial role in formal

[model selection](#) and effective validation. This methodology allows analysts to move beyond merely descriptive statistics and engage in inferential reasoning about variable importance. By comparing a rich, comprehensive model against a pared-down, nested version, researchers can formally test the necessity of including certain predictors, thereby ensuring the final chosen model is both accurate and economical.

Continuing with our basketball example, we initially established the **full model** (Model A) encompassing four predictors: minutes, height, position, and shots attempted. While Model A represents our best initial guess, theoretical insight or exploratory data analysis might suggest that 'position' and 'shots attempted' are weak predictors, potentially adding unnecessary complexity without corresponding explanatory benefit.

By implementing the **nested model** (Model B), which exclusively uses minutes played and height, we set up a direct statistical confrontation. This comparison forces an answer to the core analytical question: Does the incremental complexity introduced by the extra variables (position and shots) translate into a statistically significant enhancement in the overall model fit? Or are these added parameters simply absorbing random noise, leading to an overly complicated structure?

If the formal statistical comparison reveals that the more complex full model offers no significant advantage over the simpler nested model, then the guiding principle of parsimony demands that we select the nested model. This decision to drop unnecessary variables not only leads to a more interpretable model but also significantly mitigates the critical risk of [overfitting](#), ensuring the model performs robustly when applied to new, unseen data.

## Formal Comparison: Hypothesis Testing for Nested Models

The definitive comparison between a full model and its nested counterpart is achieved through formal [hypothesis testing](#). This rigorous statistical procedure provides the objective evidence necessary to determine whether the added complexity introduced by the extra parameters in the full model is statistically justified by a substantial improvement in fit relative to the reduced, nested model.

The comparison is formally structured around two competing hypotheses:

**The Null Hypothesis (H<sub>0</sub>):** This hypothesis asserts that the full and nested models provide an equally good fit to the data. Statistically, this is equivalent to stating that the coefficients for the supplementary variables in the full model are zero ( $\beta_{\text{extra}} = 0$ ). If the data does not provide sufficient evidence to reject H<sub>0</sub>, the simpler, **nested model is the preferred choice**, adhering to the principle of parsimony.

**The Alternative Hypothesis (H<sub>A</sub>):** This asserts that the full model fits the data significantly better

than the nested model. This implies that at least one of the coefficients associated with the additional variables is meaningfully non-zero. If  $H_0$  is rejected in favor of  $H_A$ , the **full model should be selected**, as the extra complexity is warranted by the improved fit.

Depending on the type of [regression analysis](#) employed (e.g., linear vs. generalized linear models), the primary testing mechanism varies. For models fitted using maximum likelihood estimation, the most robust test is the [Likelihood ratio test](#) (LRT). In the specific context of standard linear regression, this comparison is often performed via an F-test, which is computationally analogous to an Analysis of Variance (ANOVA) procedure comparing the residual sums of squares between the two models.

The [Likelihood ratio test](#) yields a Chi-Square test statistic, which is then used to calculate a corresponding p-value. The decision rule is straightforward: if the calculated p-value falls below a predetermined significance level (commonly set at  $\alpha = 0.05$ ), we possess strong evidence to reject the null hypothesis. This confirms that the full model provides a statistically superior fit, thereby justifying the inclusion of the additional variables. Conversely, if the p-value exceeds the threshold, we fail to reject  $H_0$ , confirming that the additional variables do not significantly enhance the model's explanatory power, and the simpler nested model is retained.

## Software Implementation of Nested Model Comparison

In applied data science and statistics, the formal comparison of nested models is streamlined and automated within specialized statistical software packages. While the foundational principles of minimizing residual error or maximizing likelihood remain constant, the specific syntaxes and dedicated functions differ across programming environments.

The standard workflow involves two steps: first, fitting both the full model and the reduced (nested) model to the dataset, and second, employing a dedicated comparison command. This command automatically computes the necessary test statistics--be it the F-statistic for linear models or the Chi-Square statistic for maximum likelihood models--based on quantifying the difference in residual sums of squares or log-likelihoods between the two model fits.

This capability is indispensable in practical machine learning and applied [statistical model](#) development. Efficiently executing the nested comparison test is the final validation step that guarantees the chosen model is not only highly accurate but also adheres to the ideal of parsimony, preventing unnecessary complexity from hindering generalization.

The following widely used statistical programming languages provide specific tools for executing the [Likelihood ratio test](#) and related comparisons:

**R Implementation:** Practitioners frequently utilize the built-in `anova()` function. When applied to

nested models generated by functions like `lm()` (for linear models) or `glm()` (for generalized linear models), `anova()` provides the requisite F-test or Chi-Square results needed for formal [hypothesis testing](#).

**Python Implementation:** Libraries such as Statsmodels offer robust support for model comparison. Nested models can often be compared using dedicated methods, or by manually calculating the test statistic based on the log-likelihood outputs obtained from the fitted models.