Understanding Significance Codes and P-Values in R for Statistical Analysis

Authored by

Mohammed looti

November 6, 2025

When performing inferential statistical tests within the **R programming environment**, such as regression analysis or ANOVA, the resulting summary tables offer essential metrics for rigorous hypothesis testing. Foremost among this output are the p-values, which provide a quantitative measure of the evidence against the null hypothesis. To supplement these precise numerical values, R automatically generates a set of visual indicators known as **significance codes**, often represented by asterisks or 'stars'.

These significance codes function as a rapid, graphical shorthand, instantly communicating whether a variable's p-value has surpassed conventional thresholds of statistical importance. For any analyst relying on R for robust statistical modeling and reporting, mastering the interpretation of these codes is not just helpful--it is fundamental to correctly communicating results.

This guide will thoroughly detail the mechanism R uses to assign these codes, demonstrate their application in common models, and discuss the necessary caveats for their responsible use in scientific reporting.

Decoding R's Significance Codes: The P-Value Thresholds

The primary function of the significance codes is to rapidly categorize the observed p-value into specific, predefined ranges. This categorization enables researchers to quickly ascertain the level of confidence associated with rejecting the null hypothesis. The system uses a hierarchy of increasingly stringent thresholds, moving from a marginal indication of significance (a decimal point) to overwhelming evidence (three asterisks).

It is paramount to recognize that these thresholds are intrinsically linked to the concept of the alpha level (α) chosen for the analysis. Although α is commonly set at 0.05, R's system provides a resolution that details where the p-value falls relative to 0.1, 0.05, 0.01, and 0.001, offering a richer context than a simple binary (significant/non-significant) assessment.

The following mapping represents the definitive system R employs internally to translate statistical evidence, as captured by the p-value, into its corresponding set of **significance codes**:

significance code p-value

****

**** (0.001, 0.01]**

*** (0.01, 0.05]**

. (0.05, 0.1]

(0.1, 1]

If a p-value is found within the range of (0.1, 1], the result is considered statistically non-significant according to these conventional metrics, and consequently, no symbol is displayed. Conversely, a p-value such as 0.00001, which signifies an extremely low probability of observing the data under the null hypothesis, falls squarely into the *** category, indicating profound statistical significance.

Setting the Context: The Alpha Level and Hypothesis Testing

Before proceeding to practical examples in R, we must firmly establish the conceptual role of the **alpha level** (α). The alpha level, typically set at 0.05 in social sciences and often lower in fields requiring greater certainty, represents the maximum acceptable risk of committing a Type I error--the mistake of rejecting a true null hypothesis. When the calculated p-value drops below this predefined α level, the result is formally classified as "statistically significant."

While R thoughtfully provides fine-grained significance codes down to the 0.001 threshold, the ultimate decision regarding the inclusion of a variable in a model or the rejection of a specific hypothesis remains the responsibility of the analyst and depends entirely on the chosen alpha level. For instance, if a researcher sets $\alpha = 0.05$, any coefficient or factor marked with a dot (.), a single star (*), two stars (**), or three stars (***) would be considered significant enough to warrant rejection of the corresponding null hypothesis.

The significance codes therefore serve as a powerful visual aid, streamlining the comparison of numerous p-values against standard benchmarks. The following sections will apply this knowledge to two widely used statistical frameworks: multiple linear regression analysis and ANOVA.

Practical Application: Interpreting Significance Codes in Regression Analysis

To demonstrate the utility of these codes, we will employ the well-known **mtcars** dataset, which is conveniently built into the R environment and frequently used for pedagogical purposes. We fit a multiple linear regression model designed to predict miles per gallon (*mpg*) based on three key predictor variables: engine horsepower (*hp*), rear axle ratio (*drat*), and vehicle weight (*wt*). The resulting output will include coefficient estimates, standard errors, and the crucial p-values for each predictor.

The commands executed below fit the model and then display the comprehensive summary output. Pay close attention to the final column of the Coefficients table, which clearly displays the critical **significance codes**:

```
#fit regression model using hp, drat, and wt as predictors
model <- lm(mpg ~ hp + drat + wt, data = mtcars)

#view model summary
summary(model)

Call:
```

lm(formula = mpg ~ hp + drat + wt, data = mtcars)

Residuals:
Min 1Q Median 3Q Max
-3.3598 -1.8374 -0.5099 0.9681 5.7078

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 29.394934 6.156303 4.775 5.13e-05 ***
hp -0.032230 0.008925 -3.611 0.001178 **
drat 1.615049 1.226983 1.316 0.198755
wt -3.227954 0.796398 -4.053 0.000364 ***

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

Residual standard error: 2.561 on 28 degrees of freedom
Multiple R-squared: 0.8369, Adjusted R-squared: 0.8194
F-statistic: 47.88 on 3 and 28 DF, p-value: 3.768e-11

Detailed Interpretation of Regression Output Codes

The column labeled `Pr(>|t|)` contains the p-values resulting from the individual t-tests for each coefficient. These values are crucial for assessing whether each predictor variable contributes significantly to the model. The symbols immediately to the right of this column are the **significance codes**, generated by R according to the thresholds defined earlier. We can now interpret the findings for each variable:

hp* (Horsepower):** The p-value for horsepower is **0.001178**. While this value is slightly above the most stringent threshold of 0.001, it falls comfortably within the range **(0.001, 0.01]**. Consequently, R assigns it the significance code of *** (two asterisks). This denotes strong evidence against the null hypothesis that horsepower has no linear association with MPG when controlling for the other variables.

***drat* (Rear Axle Ratio):** This predictor yields a p-value of **0.198755**. Because this value is substantially greater than 0.1, it falls into the statistically non-significant range **(0.1, 1]**. R displays no symbol for this term, indicating that there is insufficient evidence to confidently conclude that *drat* is a statistically significant predictor of *mpg* within the context of this specific model.

wt* (Weight):** The vehicle weight variable produces a p-value of **0.000364**. Since this value is less than 0.001 (which is equivalent to 1e-3), it is placed in the most restrictive range, . This result earns the highest code, **** (three asterisks), demonstrating overwhelming evidence that weight is a highly significant factor in predicting MPG.

Assuming the standard alpha level of α = 0.05 is chosen for reporting, we would officially conclude that both *hp* and *wt* are statistically significant predictors of *mpg*, whereas *drat* is not. The primary benefit of the significance codes is providing immediate visual confirmation of this outcome without requiring the analyst to manually compare every p-value against the established threshold.

Practical Application: Interpreting Significance Codes in ANOVA

The application and interpretation of significance codes remain identical when analyzing the output of an ANOVA model, which is typically used to test for significant differences between the means of multiple groups. In this example, we assess whether there is a statistically significant difference in mean MPG based on the number of forward gears (*gear*) present in the mtcars dataset.

We use the `aov()` function for the Analysis of Variance, treating *gear* as a categorical factor variable and *mpg* as the continuous response variable. The summary output below focuses specifically on the F-test results for the factor being examined:

#fit one-way ANOVA
model <- aov(mpg ~ gear, data = mtcars)

#view the model output
summary(model)

```
Df Sum Sq Mean Sq F value Pr(>F)
gear 1 259.7 259.75 8.995 0.0054 **
Residuals 30 866.3 28.88
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In the ANOVA context, the critical column is `Pr(>F)`, which contains the p-value derived from the F-statistic. This value determines the probability of observing the differences in group means if the null hypothesis (that all group means are equal) were true. The assigned code provides the instantaneous assessment:

gear* Factor:** The factor *gear* has a p-value of **0.0054**. This value falls within the range **(0.001, 0.01]**. Consistent with R's established methodology, it is therefore assigned a significance code of *** (two asterisks).

Utilizing the standard alpha level of α = 0.05, we would confidently conclude that the factor *gear* is statistically significant. In practical terms, this strong statistical evidence suggests that the average

MPG differs significantly across vehicles categorized by their number of gears.

Beyond the Stars: Caveats and Context in Statistical Reporting

While **significance codes** are an exceptionally efficient diagnostic tool for summarizing statistical evidence in R output, it is vital that analysts understand their inherent limitations and employ them judiciously. The presence of three stars (***) merely indicates a very small p-value; it does not, by itself, convey that the effect size is large, meaningful, or practically important.

One major caveat arises when dealing with extremely large sample sizes. In such scenarios, even trivial effect sizes--those with minimal real-world impact--can generate highly significant p-values and, consequently, earn three asterisks. This is often referred to as the distinction between statistical significance and practical significance.

To ensure robust and responsible reporting, analysts should always report the full context: the coefficient estimates (in regression) or mean differences (in ANOVA), alongside measures of effect size (such as Cohen's d, partial eta-squared, or R-squared). Significance codes are a helpful initial assessment provided by R to streamline the review of model results, but they must serve as a starting point, not the ultimate basis, for drawing scientific conclusions.