

How do I Interpret Significance Codes in R?

Authored by
stats writer

December 17, 2025

RECOMMENDED CITATION

stats writer (2025). *How do I Interpret Significance Codes in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=107685>

When conducting quantitative research or statistical analysis, researchers rely on specific indicators to determine whether observed effects are likely due to chance or represent a true underlying relationship. In the programming environment R, these indicators are often summarized using **significance codes**, providing an immediate visual cue regarding the outcome of a statistical test.

These symbols--typically stars or a dot--act as shorthand for the corresponding P-value (probability value). The P-value is the probability of observing data at least as extreme as the current data, assuming the null hypothesis is true. Therefore, understanding the relationship between these significance codes and their underlying P-values is paramount for accurate interpretation of any model output generated in R.

Contextualizing Codes in R Output

Whether you are executing a regression analysis to model relationships between variables or performing an ANOVA (Analysis of Variance) to compare group means, R's summary functions consistently provide detailed output tables. A critical component of these outputs is the column dedicated to P-values, usually labeled ``Pr(>|t|)`` for regression or ``Pr(>F)`` for ANOVA.

Adjacent to these P-values, R automatically appends the corresponding **significance codes**. These codes streamline the process of assessing **statistical significance**, allowing analysts to quickly determine which predictors or factors warrant rejection of the null hypothesis at commonly accepted alpha levels. The presence of these symbols--stars or a dot--immediately signals that the predictor variable or factor being tested has crossed a predetermined threshold of significance.

The core mapping between the P-value range and the assigned significance code is standardized across most statistical software, including R. This standardization ensures clarity when communicating results across different statistical reports. The table below illustrates the standard convention used by R:

significance code p-value

** (0.001, 0.01]

* (0.01, 0.05]

. (0.05, 0.1]

(0.1, 1]

It is essential to note that these codes are simply visual aids tied to specific hypothesis testing thresholds. Researchers must always determine their acceptable alpha level (α) beforehand. Typically, an alpha level of 0.05 is used, meaning results marked with ``*``, ``**``, or ``***``

are usually deemed statistically significant.

The Role of the Alpha Level in Significance

The decision to reject or fail to reject the null hypothesis hinges on comparing the calculated **P-value** against a predetermined significance level, known as alpha (α). The alpha level represents the maximum probability of making a Type I error--incorrectly rejecting a true null hypothesis. In many scientific fields, the conventional threshold is set at $\alpha = 0.05$.

When the P-value is less than or equal to the chosen alpha level ($P \leq \alpha$), the result is considered statistically significant, providing sufficient evidence to reject the null hypothesis. The significance codes in R directly facilitate this comparison. For instance, receiving a `***` code means the P-value is less than 0.01, which is substantially smaller than the standard $\alpha=0.05$ threshold, indicating strong evidence against the null hypothesis.

However, analysts should exercise caution when interpreting the boundary case marked by the dot (`.`), corresponding to P-values between 0.05 and 0.1. While this range is often referred to as marginally or tententially significant, it does not meet the strict $\alpha=0.05$ standard. Researchers must decide based on their field's conventions whether to treat results falling in this marginal range as meaningful.

Example: Interpreting Significance Codes in Regression Analysis

To demonstrate the practical application of these significance codes, we will examine the output of a multiple linear regression analysis performed using R. Regression models seek to quantify the relationship between one dependent variable (response) and multiple independent variables (predictors). For this illustration, we utilize the famous built-in R dataset, **mtcars**, which contains specifications and performance metrics for 32 automobiles.

We model miles per gallon (*mpg*, the response variable) as a function of horsepower (*hp*), rear axle ratio (*drat*), and vehicle weight (*wt*). The following R code snippet executes the model fitting using the `lm()` function and displays the detailed statistical summary using `summary()`, which includes the coefficients table containing our target 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

Analysis of Regression Predictor Significance

The coefficients table provides the necessary statistics to evaluate each predictor's unique contribution to the model. We focus specifically on the `Pr(>|t|)` column, which reports the **P-value** for the test associated with the null hypothesis that the coefficient is zero. The appended significance codes clearly map these probabilities to common thresholds:

Horsepower (*hp*): The corresponding P-value is **0.001178**. This value falls immediately above the 0.001 cutoff but remains well below the 0.01 threshold. Consequently, R assigns the significance code of `**`. This indicates a highly significant relationship, providing strong evidence that horsepower is a non-zero predictor of MPG.

Rear Axle Ratio (*drat*): The P-value for *drat* is **0.198755**. Since this probability is greater than 0.1, it falls into the broadest range defined by R and receives no specific symbol (indicated by a space ` `). This large P-value suggests that, holding the other variables constant, *drat* is not a **statistically significant** predictor of MPG in this particular model.

Weight (*wt*): The P-value for vehicle weight is **0.000364**. This value is extremely small, falling squarely within the most stringent significance band, which is $P \leq 0.001$. As a result, R assigns the code `***`. This strong result suggests a highly robust, non-zero relationship between vehicle weight and MPG.

Based on the standard $\alpha = 0.05$ threshold for **statistical significance**, we would formally conclude that both *hp* and *wt* are statistically significant predictors within this multiple linear

regression model. Conversely, we must fail to reject the null hypothesis for *drat*, meaning its observed relationship with *mpg* could plausibly be due to random chance.

Example: Interpreting Significance Codes in ANOVA

The methodology for interpreting significance codes remains consistent across different modeling approaches, including ANOVA. ANOVA is specifically designed to test for differences in means across two or more groups categorized by a factor variable. Here, we utilize a one-way ANOVA to determine if the mean miles per gallon (*mpg*) differs significantly based on the number of forward gears (*gear*), treating *gear* as a categorical factor.

In R, the `aov()` function is typically used for this analysis. The summary output displays an ANOVA table that uses the F-test statistic to evaluate the overall effect of the factor variable. The P-value associated with this F-test determines the factor's significance, which R again summarizes using the familiar code system.

#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
```

Conclusion from the ANOVA Summary

In the ANOVA summary table, we focus on the row corresponding to the factor variable, *gear*. The crucial metric here is the **P-value**, listed under the `Pr(>F)` column, which tests the null hypothesis that the mean MPG is identical across all groups defined by the *gear* variable.

Factor *gear*: The calculated P-value is **0.0054**. This value is clearly smaller than 0.01 but larger than 0.001. According to the standard significance codes defined by R, this range corresponds to the `**` symbol.

Given that the P-value (0.0054) is substantially less than the customary alpha level of $\alpha = 0.05$, we conclude that the factor *gear* is highly **statistically significant**. This finding allows us to confidently reject the null hypothesis. Substantively, this means there is robust statistical evidence

demonstrating that the average miles per gallon differs significantly among cars depending on the number of forward gears they possess.

Understanding these significance codes in R is essential for efficiently reading statistical output. While the codes provide a rapid visual assessment, researchers should always reference the exact P-value to make precise judgments, especially when results fall near the 0.05 or 0.1 thresholds.

ARABPSYCHOLOGY.COM