

How to Easily Perform a t-Test for Regression Slope in R

Authored by
stats writer

November 21, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform a t-Test for Regression Slope in R*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=98840>

Determining the statistical significance of the relationship between a predictor variable and a response variable is a core task in statistical analysis. When using Linear Regression, this determination hinges on assessing the Slope Coefficient of the estimated regression line. If the slope is statistically different from zero, we can confidently conclude that the predictor has a measurable effect on the outcome. The most common method used to formally test this relationship is the t-test for the slope.

While some statistical environments require the manual calculation of the t-statistic or the use of specialized functions, the statistical programming language R simplifies this process significantly. In R, fitting a linear model using the `lm()` function automatically generates all the necessary outputs, including the t-statistic and corresponding P-value, allowing for immediate interpretation of the slope's significance. Understanding how to precisely interpret the results from the `summary()` function is therefore crucial for any data analyst working with regression models in R.

This detailed guide will walk through the theoretical framework of the t-test for the slope, explain the structure of the regression model, and demonstrate step-by-step how to perform and interpret this critical statistical test using R's powerful built-in functions. We will focus on ensuring a clear, structured understanding of why and how the slope coefficient is tested against the null hypothesis.

Theoretical Foundation of the Regression Slope Test

Whenever we perform simple Linear Regression, we are attempting to model the relationship between a single independent variable (X) and a dependent variable (Y). This analysis yields an estimated sample regression equation, which mathematically describes the best-fit line through the observed data points. This line is expressed as:

$$\hat{Y} = b_0 + b_1x$$

In this equation, \hat{Y} represents the predicted value of the dependent variable, b_0 is the y-intercept (the predicted value of Y when X is zero), and b_1 is the Slope Coefficient. The slope, b_1 , is the critical parameter, as it quantifies the average change in Y associated with a one-unit increase in X. If the true slope in the population (often denoted as β) is zero, then X has no linear relationship with Y, meaning the predictor variable holds no predictive power in the model.

Our primary statistical goal is to determine if the calculated sample slope, b_1 , provides sufficient evidence to conclude that the true population slope (β) is non-zero. This determination requires formal Hypothesis Testing, which assesses the likelihood of observing our calculated slope if, in reality, no underlying relationship existed in the broader population.

Formulating the Null and Alternative Hypotheses

The t-test for the Slope Coefficient is structured around two competing statements concerning the population parameter β --the true, underlying slope of the relationship.

The Null Hypothesis (H_0)

The null hypothesis always asserts a lack of effect or relationship. In the context of linear regression, H_0 states that the true slope is zero. If we fail to reject H_0 , we conclude that the observed sample relationship is likely a result of random sampling variation or chance.

$H_0: \beta = 0$ (The population slope is zero; the predictor variable has no statistically significant linear effect on the response variable.)

The Alternative Hypothesis (H_1 or H_a)

The alternative hypothesis asserts that a statistically significant relationship exists. This means the true slope is not zero. While directional tests (one-tailed) exist, the standard procedure performed by R's `lm()` function is a two-tailed test, checking for significance in either the positive or negative direction.

$H_1: \beta \neq 0$ (The population slope is not zero; a statistically significant linear relationship exists.)

The decision to reject H_0 is fundamentally driven by the probability (P-value) of obtaining our sample data if the null hypothesis were true. If this probability is very low (typically less than 5%), we reject the null hypothesis in favor of the alternative.

Deriving the T-Statistic and Understanding Standard Error

To quantify how far our estimated sample slope **b** lies from the null value of zero, we calculate the t-statistic. The t-statistic measures this deviation in units of Standard Error. The general formula for the t-statistic of the slope is:

$$t = (b - \beta_{\text{hypothesized}}) / \text{se}(b)$$

Since we test against the null hypothesis that β is zero, the formula simplifies significantly:

$$t = b / \text{se}(b)$$

Where:

b represents the estimated Slope Coefficient (the value in the "Estimate" column of the R output).

se(b) represents the Standard Error of the slope estimate (the value in the "Std. Error" column).

The **Standard Error (se(b₁))** is a critical measure of precision. It estimates the variability of the sample slope estimate across hypothetical repetitions of the sampling process. A small standard error suggests that our estimate **b₁** is highly reliable and precise. A large t-statistic (far from zero) indicates that the observed slope is many standard errors away from zero, making it statistically unlikely to have occurred if H₀ were true. This test statistic follows a t-distribution with **n-2** degrees of freedom for simple linear regression.

Practical Implementation: Linear Regression in R

R automates the complex calculations involved in the t-test for the slope through the `lm()` function. We do not need to manually calculate the t-statistic or the P-value; we only need to invoke the model summary.

We will use the following example data frame, containing the hours studied and the final exam scores for 12 students:

Step 1: Data Setup

The dataset includes 12 observations (n=12), establishing the degrees of freedom for the t-test as $12 - 2 = 10$. We hypothesize that 'hours' predicts 'score'.

#create data frame

```
df <- data.frame(hours=c(1, 1, 2, 2, 3, 4, 5, 5, 5, 6, 6, 8),
score=c(65, 67, 78, 75, 73, 84, 80, 76, 89, 91, 83, 82))
```

```
#view data frame
```

```
df
```

```
hours score
```

```
1 1 65
```

```
2 1 67
```

```
3 2 78
```

```
4 2 75
```

```
5 3 73
```

```
6 4 84
```

```
7 5 80
```

```
8 5 76
```

```
9 5 89
```

```
10 6 91
```

```
11 6 83
```

```
12 8 82
```

Step 2: Fitting the Model and Generating Summary

We use the `lm()` function to fit the model and then apply `summary()` to obtain the full regression output, which contains the results of the t-test for both the intercept and the slope:

```
#fit simple linear regression model
```

```
fit <- lm(score ~ hours, data=df)
```

```
#view model summary
```

```
summary(fit)
```

Call:

```
lm(formula = score ~ hours, data = df)
```

Residuals:

```
Min 1Q Median 3Q Max
```

```
-7.398 -3.926 -1.139 4.972 7.713
```

Coefficients:

```
Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 67.7685 3.3757 20.075 2.07e-09 ***
```

```
hours 2.7037 0.7456 3.626 0.00464 **
```

```
---
```

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

Residual standard error: 5.479 on 10 degrees of freedom

Multiple R-squared: 0.568, Adjusted R-squared: 0.5248

F-statistic: 13.15 on 1 and 10 DF, p-value: 0.004641

Interpreting the R Output for Slope Significance

The interpretation centers on the "Coefficients" table within the `summary(fit)` output. This table provides all the necessary statistics to test the null hypothesis for the Slope Coefficient ('hours').

Focusing on the Slope Coefficient ('hours')

The estimated regression equation is: **Estimated Score = 67.7685 + 2.7037(Hours Studied)**. We focus on the row corresponding to the 'hours' variable:

Estimate (b?): 2.7037. This is the calculated slope, indicating a positive relationship.

Std. Error (se(b?)): 0.7456. This is the measure of the precision of our slope estimate.

t value (t): 3.626. This is the calculated t-statistic. It confirms that the observed slope (2.7037) is

approximately 3.626 standard errors away from zero.

Pr(>|t|) (P-value): 0.00464. This is the probability of observing a t-statistic as extreme as 3.626 (or more extreme) if the null hypothesis ($\beta = 0$) were true.

Making the Statistical Decision

To conclude the t-test, we compare the calculated P-value to the chosen significance level (α), which is almost universally set at 0.05. The decision rule is simple: If $P\text{-value} < \alpha$, reject H_0 .

P-value = 0.00464

Significance Level (α) = 0.05

Since 0.00464 is substantially less than 0.05, we reject the null hypothesis ($H_0: \beta = 0$). We conclude with strong evidence that the Slope Coefficient is statistically significant, meaning that studying hours genuinely influences the final exam score.

The Critical Role of Precision and Confidence Intervals

The precision provided by the Standard Error is fundamental to the t-test's validity. If the standard error were larger--perhaps if the data were highly scattered or the sample size was small--the t-statistic would decrease, leading to a higher P-value. Thus, a robust finding requires not only a large slope estimate but also a relatively small standard error.

Furthermore, the standard error allows us to construct a confidence interval for the true population slope (β). If the 95% confidence interval for the slope does not contain the value zero, the coefficient is deemed statistically significant at the $\alpha = 0.05$ level. This confidence interval calculation provides an equivalent decision pathway to the t-test, offering a range of plausible values for the true relationship strength. For our example, since we rejected H_0 , the 95% confidence interval for the hours coefficient (2.7037) will definitely exclude zero, reinforcing the conclusion derived from the Hypothesis Testing.

Conclusion and Final Summary

The execution of a t-test for the slope of a Linear Regression line in R is streamlined by the output generated from the `summary(lm())` function. This process is essential for transitioning from a descriptive observation (the sample slope b) to an inferential conclusion about the true population relationship (β).

By focusing on the t value and, most importantly, the corresponding **Pr(>|t|)** column in the coefficient summary, analysts can quickly determine if the predictor variable has a genuine, statistically significant linear impact on the response. The successful rejection of the null

hypothesis confirms that the estimated slope is not merely a random fluctuation but represents a meaningful effect worthy of interpretation and reporting.

ARABPSYCHOLOGY.COM