

Testing the Significance of a Regression Slope: A Step-by-Step Guide

Authored by
Mohammed loot

November 9, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Testing the Significance of a Regression Slope: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=14502>

Establishing the Foundation: Analyzing the Sample Dataset

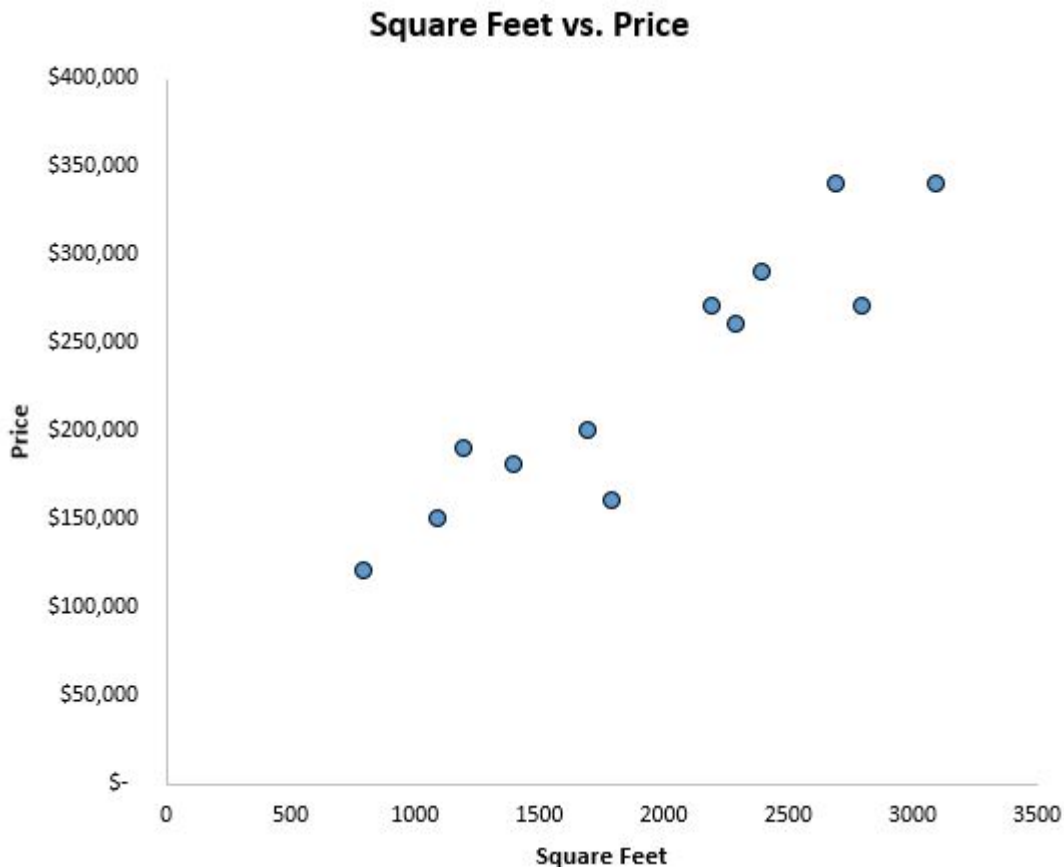
Before diving into the mechanics of hypothesis testing, we must first establish a clear context. Our objective is to determine the [statistical significance](#) of a [regression slope](#)--a fundamental step in validating predictive models. To illustrate this crucial process, we utilize a focused sample dataset. This dataset comprises observations from 12 distinct houses, meticulously recording two key variables: the house size (measured in square feet) and its final sale price.

Square Feet	Price
800	\$120,000
1400	\$180,000
1200	\$190,000
2300	\$260,000
1800	\$160,000
2700	\$340,000
2400	\$290,000
1100	\$150,000
1700	\$200,000
3100	\$340,000
2800	\$270,000
2200	\$270,000

The central question driving this analysis is whether a meaningful, non-random relationship exists between property size and market value. If we confirm that this relationship is statistically significant, it means that house size is a robust and reliable predictor of price, elevating our conclusion above speculative observation. This foundational dataset provides the essential data points required for our subsequent statistical investigation and modeling efforts, setting the stage for proving a true relationship exists in the population.

Initial Assessment: Visualizing the Relationship with a Scatterplot

Prior to fitting any statistical model, it is standard practice to visualize the data. This crucial step helps confirm the linearity assumption required for a [simple linear regression](#). We generate a [scatterplot](#) where the **predictor variable** (square feet) is plotted on the horizontal (x) axis, and the **response variable** (price) is plotted on the vertical (y) axis. This graphical method instantly reveals the underlying structure, direction, and strength of the association between the two variables.



Upon inspection, the scatterplot demonstrates a compelling **positive linear correlation**. The data points cluster tightly around what appears to be an upward-sloping line, indicating that as the size of the house increases, the sale price predictably rises. This strong visual evidence confirms that applying a linear model is highly appropriate for quantifying the observed trend. However, while visualization suggests a relationship, formal statistical testing is necessary to provide rigorous, quantifiable proof of significance that holds up under scrutiny.

Quantifying the Trend: Executing the Regression Model

To transition from visual confirmation to numerical quantification, we execute a [simple linear regression](#) analysis. This statistical procedure formalizes the relationship by defining the line of best fit that minimizes the residuals (the differences between the observed data points and the predicted values). In our model, square footage acts as the independent variable (x), aiming to predict the dependent variable, price (y). The resulting output from this analysis is standardized across most statistical platforms, providing the core parameters needed for interpreting the relationship and testing its reliability.

The critical results of the regression analysis are summarized in the coefficient table below. This table contains the essential estimates and measures of uncertainty, such as the [standard error](#),

which are indispensable for conducting our formal [hypothesis test](#) for the slope.

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	47588.70	23838.22	2.00	0.07
Square Feet	93.57	11.45	8.17	0.00

This regression output table forms the backbone of our subsequent analysis. It not only yields the calculated values for the y-intercept and the slope but also provides crucial statistics, including the t-statistic and the associated [p-value](#). These metrics are the direct indicators we will use to evaluate whether the predictive power exerted by square footage is truly **statistically significant**, moving beyond mere estimation to robust inference.

Defining the Line of Best Fit and Interpreting Coefficients

The output from the regression analysis allows us to construct the specific mathematical equation for the line of best fit. This equation is the foundation of our predictive model and is universally expressed in the form:

$$? = b_0 + b_1x$$

In this formula, ? represents the predicted value (house price), x is the predictor variable (square footage), b₀ is the estimated y-intercept, and b₁ is the estimated [regression coefficient](#), or slope. Extracting the numerical values directly from our model output yields the following parameters:

The Y-Intercept (b₀) is calculated as **47588.70**.

The Slope (b₁) for the variable *Square Feet* is **93.57**.

Substituting these coefficients, our tailored predictive model becomes **? = 47588.70 + 93.57x**. The interpretation of the slope (b₁) is critical for practical understanding, as it defines the marginal effect of the predictor variable on the response:

Practical Interpretation: For every one-unit increase in the size of a house (one square foot), the model predicts an average increase in the house price of **\$93.57**.

While the slope estimate of 93.57 suggests a clear positive relationship, this value is based solely on our sample data. The paramount statistical task remains: determining if this sample slope is sufficiently far from zero to conclude that the true population slope (B₁) is non-zero. If the true slope were zero, the relationship would be considered nonexistent or spurious. We can rigorously test this crucial significance using one of two methods: the [confidence interval](#) approach or the formal [hypothesis test](#).

Method 1: Assessing Significance Using a Confidence Interval

One highly intuitive method for testing the significance of the population slope (B_1) is constructing a [confidence interval](#). This interval defines a plausible range of values where the true population parameter is expected to reside, based on our sample data. The key rule for significance is simple: if the calculated range for the slope entirely excludes the value zero, we can confidently reject the notion that no relationship exists.

The calculation of the confidence interval relies on three components: the estimated slope (b_1), the critical t-value (which is determined by the chosen confidence level and degrees of freedom), and the [standard error](#) of the slope estimate. The formula encapsulates this relationship:

$$\text{Confidence Interval} = b_1 \pm (t_{1-\alpha/2, n-2}) * (\text{Standard Error of } b_1)$$

For our analysis, we aim to construct a 95% confidence interval (meaning $\alpha = 0.05$). We draw the necessary input parameters directly from the regression output and statistical tables:

Estimated Slope (b_1): 93.57.

Critical t-value: Given 12 observations, the degrees of freedom ($n-2$) is 10. For a two-tailed 95% interval, the critical t-value is **2.228**.

Standard Error of b_1 : This measure of variability is 11.45, as reported in the regression summary table.

Substituting these values into the formula allows us to calculate the margin of error and define the bounds of the 95% confidence interval for B_1 :

$$93.57 \pm (2.228) * (11.45) = \mathbf{(68.06, 119.08)}$$

The interpretation of this result is straightforward: we are 95% certain that the true average increase in price per additional square foot of space falls somewhere between \$68.06 and \$119.08. Critically, because this entire interval is positive and definitively **does not contain zero**, we have strong [statistical evidence](#) to conclude that the relationship between square footage and price is **statistically significant** at the 95% confidence level.

Method 2: The Formal Hypothesis Test for Regression Slope

While the confidence interval offers a range of plausible values, the formal [hypothesis test](#) provides a structured, multi-step statistical proof that leads to the exact same conclusion regarding significance. This method is often preferred for academic and formal reporting purposes as it directly addresses the null hypothesis.

Step 1: Define the Hypotheses. The test is designed to determine if the population slope (B_1) is

non-zero. We set up the hypotheses as follows:

Null Hypothesis (H_0): $B_1 = 0$. This assumes that square footage has absolutely no linear effect on price.

Alternative Hypothesis (H_a): $B_1 \neq 0$. This is our research hypothesis, suggesting a genuine linear relationship exists.

Step 2: Set the Significance Level. To maintain consistency with the confidence interval method, we establish a significance level (α , ∞) of **0.05**. This means we are willing to accept a 5% chance of incorrectly rejecting the null hypothesis (Type I error).

Step 3: Calculate the Test Statistic and P-value. The t-test statistic measures how many [standard errors](#) the estimated slope (b_1) is away from the hypothesized value of zero. The formula is $t = b_1 / SE(b_1)$. This value is often provided directly in the regression output:

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	49220.57	29989.65	1.64	0.14
Square Feet	92.89	13.88	6.69	0.00

Using our regression coefficients ($b_1 = 93.57$ and $SE = 11.45$), the calculated test statistic is $t = 93.57 / 11.45 \approx 8.17$. With 10 degrees of freedom, this exceptionally high t-score corresponds to a two-tailed [p-value](#) that is extremely small, typically reported as $p < 0.001$ or rounded to **0.000**.

Step 4: Make the Decision. The decision rule dictates that if the p-value is less than the significance level (∞), we reject H_0 . Since our p-value (0.000) is far smaller than ∞ (0.05), we emphatically **reject the null hypothesis**.

Step 5: Conclude the Finding. By rejecting the null hypothesis, we confirm that there is overwhelming [statistical evidence](#) to conclude that the true average increase in price for each additional square foot is not zero. The slope of the [regression line](#) is therefore **statistically significant**, definitively establishing square footage as a statistically valid and powerful predictor of house price in this market sample.