What is the meaning and significance of the standard error of a regression slope?

Authored by

stats writer

May 12, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the meaning and significance of the standard error of a regression slope?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=143942

The standard error of a regression slope is a measure of the variability or uncertainty associated with the estimated slope of a regression line. It measures the average difference between the actual data points and the predicted values by the regression line. This statistic is significant because it helps to assess the accuracy and reliability of the estimated slope, and it is used to determine the statistical significance of the relationship between the independent and dependent variables. A smaller standard error indicates a more precise estimate of the slope, while a larger standard error indicates a less reliable estimate. Overall, understanding the standard error of a regression slope is crucial in evaluating the strength and validity of a regression analysis.

Understanding the Standard Error of a Regression Slope

The standard error of a regression slope is a way to measure the "uncertainty" in the estimate of a regression slope.

It is calculated as:

$$s(b_1) = \sqrt{\frac{1}{n-2} * \frac{\sum(y_i - \hat{y}_i)^2}{\sum(x_i - \bar{x})^2}}$$

where:

n: total sample sizeyi: actual value of response variable?i: predicted value of response variablexi: actual value of predictor variablex?: mean value of predictor variable

The smaller the standard error, the lower the variability around the coefficient estimate for the regression slope.

The standard error of the regression slope will be displayed in a "standard error" column in the regression output of most statistical software:

Regression Statistics	
Multiple R	0.939
R Square	0.882
Adjusted R Square	0.877
Standard Error	3.246
Observations	25

ANOVA

	df	*SS*	*MS*	*F*	*p*
Regression	1	1811.489601	1811.4896	171.91744	3.6975E-12
Residual	23	242.3503989	10.536974		
Total	24	2053.84			

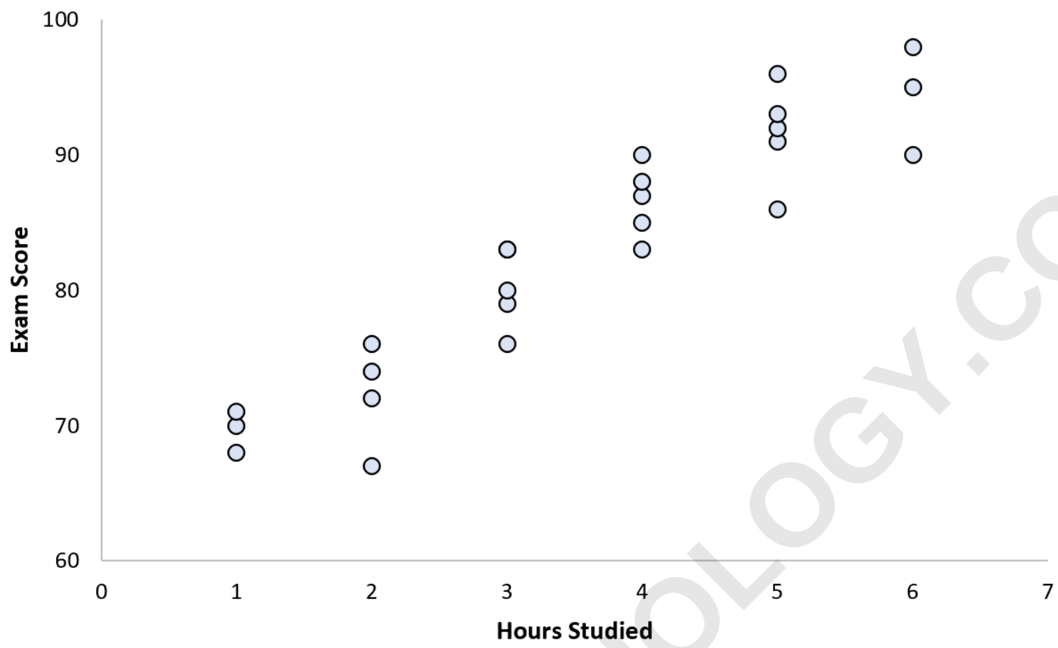
	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	63.385	1.625	39.002	0.000
hours	5.487	**0.419**	13.112	0.000

The following examples show how to interpret the standard error of a regression slope in two different scenarios.

Example 1: Interpreting a Small Standard Error of a Regression Slope

Suppose a professor wants to understand the relationship between the number of hours studied and the final exam score received for students in his class.

He collects data for 25 students and creates the following scatterplot:



There is a clear positive association between the two variables. As hours studied increases, the exam score increases at a fairly predictable rate.

He then fits a simple linear regression model using hours studied as the predictor variable and final exam score as the response variable.

The following table shows the results of the regression:

Regression Statistics	
Multiple R	0.939
R Square	0.882
Adjusted R Square	0.877
Standard Error	3.246
Observations	25

ANOVA

	df	*SS*	*MS*	*F*	*p*
Regression	1	1811.489601	1811.4896	171.91744	3.6975E-12
Residual	23	242.3503989	10.536974		
Total	24	2053.84			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	63.385	1.625	39.002	0.000
hours	5.487	0.419	13.112	0.000

The standard error is 0.419, which is a measure of the variability around this estimate for the regression slope.

We can use this value to calculate the t-statistic for the predictor variable 'hours studied':

t-statistic = coefficient estimate / standard errort-statistic = 5.487 / .419t-statistic = 13.112

The p-value that corresponds to this test statistic is 0.000, which indicates that 'hours studied' has a statistically significant relationship with final exam score.

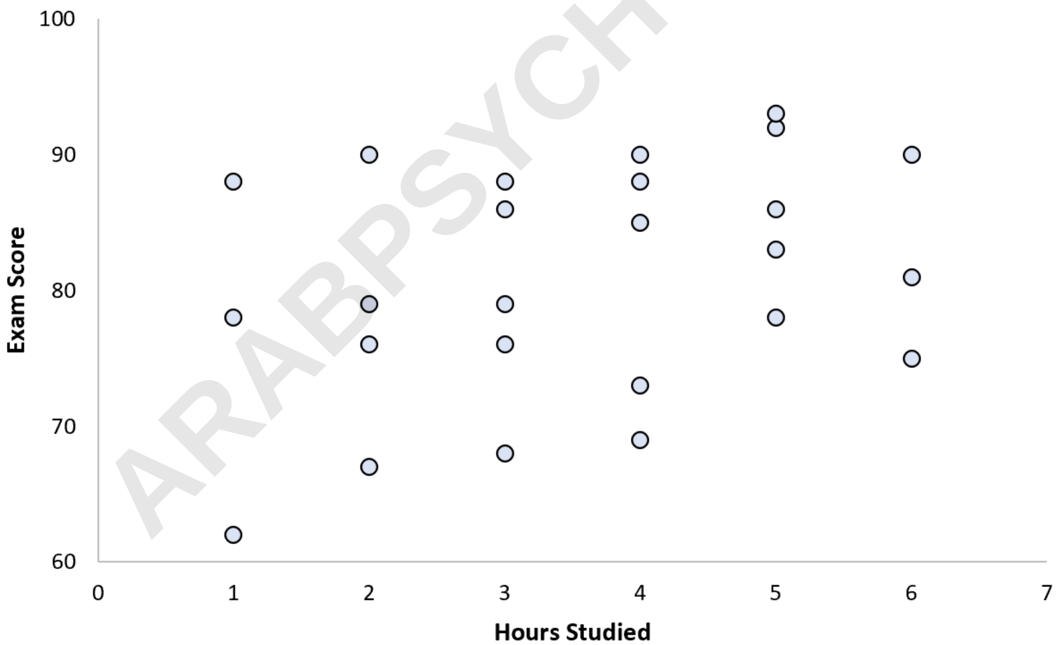
Since the standard error of the regression slope was

small relative to the coefficient estimate of the regression slope, the predictor variable was statistically significant.

Example 2: Interpreting a Large Standard Error of a Regression Slope

Suppose a different professor wants to understand the relationship between the number of hours studied and the final exam score received for students in her class.

She collects data for 25 students and creates the following scatterplot:



There appears to be slight positive association between the two variables. As hours studied increases, the exam

score generally increases but not at a predictable rate.

Suppose the professor then fits a simple linear regression model using hours studied as the predictor variable and final exam score as the response variable.

The following table shows the results of the regression:

Regression Statistics	
Multiple R	0.330353
R Square	0.1091331
Adjusted R Square	0.0703997
Standard Error	8.279977
Observations	25

ANOVA

	df	*SS*	*MS*	*F*	*p*
Regression	1	193.1655585	193.16556	2.8175487	0.10677595
Residual	23	1576.834441	68.558019		
Total	24	1770			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	74.420878	4.145490151	17.952251	5.023E-15
hours	1.7918883	1.067518025	1.6785555	0.106776

The coefficient for the predictor variable 'hours studied' is 1.7919. This tells us that each additional hour studied is associated with an average increase of 1.7919 in exam score.

The standard error is 1.0675, which is a measure of the variability around this estimate for the regression slope.

We can use this value to calculate the t-statistic for the predictor variable 'hours studied':

t-statistic = coefficient estimate / standard errort-statistic = 1.7919 / 1.0675t-statistic = 1.678

The p-value that corresponds to this test statistic is 0.107. Since this p-value is not less than .05, this indicates that 'hours studied' does not have a statistically significant relationship with final exam score.

Since the standard error of the regression slope was large relative to the coefficient estimate of the regression slope, the predictor variable was *not* statistically significant.