What is the difference between Multiple R and R-Squared?

Authored by

stats writer

April 25, 2024

Multiple R and R-Squared are two commonly used statistical measures in regression analysis. Both of these measures are used to assess the strength and direction of the relationship between two variables. However, there are some key differences between the two.

Multiple R, also known as the multiple correlation coefficient, measures the overall strength of the linear relationship between the dependent variable and all the independent variables in a regression model. It ranges from -1 to 1, with a higher value indicating a stronger relationship. Multiple R does not take into account the number of independent variables in the model.

On the other hand, R-Squared, also known as the coefficient of determination, measures the proportion of variation in the dependent variable that can be explained by the independent variables in the model. It ranges from 0 to 1, with a higher value indicating a better fit of the model. Unlike Multiple R, R-Squared takes into account the number of independent variables in the model and penalizes for adding unnecessary variables.

In summary, Multiple R measures the strength of the relationship while R-Squared measures the goodness of fit of the model. Both measures are important in understanding and interpreting the results of a regression analysis.

Multiple R vs. R-Squared: What's the Difference?

When you fit a using most statistical software, you'll often notice the following two values in the output:

Multiple R: The multiple correlation coefficient between three or more variables.

R-Squared: This is calculated as (Multiple R)2 and it represents the proportion of the variance in the of a regression model that can be explained by the predictor variables. This value ranges from 0 to 1.

In practice, we're often interested in the R-squared

value because it tells us how useful the predictor variables are at predicting the value of the response variable.

However, each time we add a new predictor variable to the model the R-squared is guaranteed to increase even if the predictor variable isn't useful.

The adjusted R-squared is a modified version of R-squared that adjusts for the number of predictors in a regression model. It is calculated as:

Adjusted R2 = 1 -

where:

R2: The R2 of the modeln: The number of observationsk: The number of predictor variables

Since R-squared always increases as you add more predictors to a model, adjusted R-squared can serve as a metric that tells you how useful a model is, *adjusted for the number of predictors in a model*.

To gain a better understanding of each of these terms, consider the following example.

Example: Multiple R, R-Squared, & Adjusted R-Squared

Suppose we have the following dataset that contains the following three variables for 12 different students:

Study Hours	Current Grade	Exam Score
1	65	58
1	78	61
2	76	62
2	76	65
1	79	65
2	80	68
2	81	72
3	84	74
3	88	78
4	85	85
4	96	90
5	90	95

Suppose we fit a multiple linear regression model using *Study Hours* and *Current Grade* as the predictor variables and *Exam Score* as the response variable and get the following output:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.978
R Square	0.956
Adjusted R Square	0.946
Standard Error	2.790
Observations	12

ANOVA

	df	SS	MS	F	p-value
Regression	2	1516.192	758.096	97.388	0.000
Residual	9	70.058	7.784		
Total	11	1586.250			

	Coefficients	Standard Error	t Stat	P-value
Intercept	17.175	12.556	1.368	0.205
Study Hours	6.384	1.087	5.874	0.000
Current Grade	0.486	0.179	2.709	0.024

We can observe the values for the following three metrics:

R Square: 0.956. This is calculated as (Multiple R)2 = (0.978)2 = 0.956. This tells us that 95.6% of the variation in exam scores can be explained by the number of hours spent studying by the student and their current grade in the course.

Adjusted R-Square: 0.946. This is calculated as:

Adjusted R2= 1 - = 1 - = 0.946.

This represents the R-squared value, *adjusted for the number of predictor variables in the model*.

This metric would be useful if we, say, fit another regression model with 10 predictors and found that the Adjusted R-squared of that model was 0.88. This would indicate that the regression model with just two predictors is better because it has a higher adjusted R-squared value.