How do I perform a Partial F-Test in Excel?

Authored by

stats writer

April 28, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I perform a Partial F-Test in Excel?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=140672

A Partial F-Test in Excel is a statistical tool used to compare the significance of two regression models. It helps determine if adding a new variable to a model significantly improves its predictive power. This test can be performed by using the built-in Analysis ToolPak add-in in Excel. First, select the data and go to the Data tab, then click on "Data Analysis" and choose "Anova: Two-Factor with Replication" from the list. Next, select the variables for the "Model" and "Residual" fields and click "OK" to generate the results, including the F-statistic and p-value. A low p-value indicates that the new variable has a significant impact on the model's performance. This test is useful in evaluating the necessity of adding new variables to a model and making informed decisions in data analysis.

Perform a Partial F-Test in Excel

A is used to determine whether or not there is a statistically significant difference between a regression model and some nested version of the same model.

A *nested* model is simply one that contains a subset of the predictor variables in the overall regression model.

For example, suppose we have the following regression model with four predictor variables:

Y = β0 + β1x1 + β2x2 + β3x3 + β4x4 + ε

One example of a nested model would be the following model with only two of the original predictor variables:

Y = β0 + β1x1 + β2x2 + ε

To determine if these two models are significantly different, we can perform a partial F-test, which calculates the following F test statistic:

F = ((RSSreduced - RSSfull)/p) / (RSSfull/n-k)

where:

RSSreduced: The residual sum of squares of the reduced (i.e. "nested") model.RSSfull: The residual sum of squares of the full model.p: The number of predictors removed from the full model.n: The total observations in the dataset.k: The number of coefficients (including the intercept) in the full model.

This test uses the following null and alternative :

H0: All coefficients removed from the full model are zero.

HA: At least one of the coefficients removed from the full model is non-zero.

If the corresponding to the F test-statistic is below a certain significance level (e.g. 0.05), then we can reject the null hypothesis and conclude that at least one of the coefficients removed from the full model is significant.

The following example shows how to perform a partial F-test in Excel.

Example: Partial F-Test in Excel

Suppose we have the following dataset in Excel:

	A	B	C	D	E	F	G
1	x1	x2	x3	x4	y		
2	1	15	23	19	18		
3	3	14	25	19	20		
4	3	15	24	17	24		
5	4	12	24	17	24		
6	4	10	14	15	23		
7	5	9	9	18	26		
8	7	6	17	15	29		
9	8	7	17	13	30		
10	8	5	4	19	37		
11	9	4	13	8	34		
12	10	4	6	8	38		
13	12	3	8	17	40		
14							
15							
16							
17							
18							
19							
20							
21							
22							

Suppose we would like to determine if there is a difference between the following two regression models:

Full Model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$

Reduced Model: y = β0 + β1x1 + β2x2

We can proceed to perform for each model to get the following output:

G	H	I	J	K	L	M
Full Model	ANOVA					
		df	*SS*	*MS*	*F*	*ignificance F*
	Regression	4	561.6047	140.4012	38.82785	7.24E-05
	Residual	7	25.31194	3.615992		
	Total	11	586.9167			
		Coefficients	*andard Errc*	*t Stat*	*P-value*	*Lower 95% l*
	Intercept	14.34009	7.565093	1.89556	0.099856	-3.54851
	x1	2.290687	0.59909	3.82361	0.006512	0.874064
	x2	0.516173	0.567794	0.909086	0.393528	-0.82645
	x3	-0.29052	0.153701	-1.89018	0.100649	-0.65397
	x4	0.006391	0.189974	0.033642	0.974102	-0.44283
	SUMMARY OUTPUT					
	Regression Statistics					
	Multiple R	0.965112				
	R Square	0.931441				
	Adjusted R	0.916206				
	Standard E	2.114459				
	Observatio	12				
Reduced Model	ANOVA					
		df	*SS*	*MS*	*F*	*ignificance F*
	Regression	2	546.6782	273.3391	61.13686	5.78E-06
	Residual	9	40.23844	4.470938		
	Total	11	586.9167			
		Coefficients	*andard Errc*	*t Stat*	*P-value*	*Lower 95% l*
	Intercept	17.50073	7.99721	2.188354	0.056401	-0.59022
	x1	1.965388	0.640132	3.070286	0.013348	0.517309
	x2	-0.11969	0.475321	-0.2518	0.806849	-1.19494

We can then use the following formula to calculate the F

test-statistic for the partial F-test:

G	H	I	J	K	L	M	N	O	P	Q	R
Full Model	ANOVA								**F test statistic**	2.064	=(J12-J34)/2/K13

Full Model — ANOVA

	df	SS	MS	F	ignificance F
Regression	4	561.6047	140.4012	38.82785	7.24E-05
Residual	7	25.31194	3.615992		
Total	11	586.9167			

	Coefficients	andard Errc	t Stat	P-value	Lower 95%	Upper 95%	ower 95.0%	Upper 95.0%
Intercept	14.34009	7.565093	1.89556	0.099856	-3.54851	32.22869	-3.54851	32.22869351
x1	2.290687	0.59909	3.82361	0.006512	0.874064	3.70731	0.874064	3.707309921
x2	0.516173	0.567794	0.909086	0.393528	-0.82645	1.858791	-0.82645	1.858791347
x3	-0.29052	0.153701	-1.89018	0.100649	-0.65397	0.072923	-0.65397	0.072923129
x4	0.006391	0.189974	0.033642	0.974102	-0.44283	0.455609	-0.44283	0.45560912

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.965112
R Square	0.931441
Adjusted R	0.916206
Standard E	2.114459
Observatio	12

Reduced Model ANOVA

	df	SS	MS	F	ignificance F
Regression	2	546.6782	273.3391	61.13686	5.78E-06
Residual	9	40.23844	4.470938		
Total	11	586.9167			

The test statistic turns out to be 2.064.

We can then use the following formula to calculate the corresponding p-value:

G	H	I	J	K	L	M	N	O	P	Q	R	S	T
Full Model	ANOVA								**F test statistic**	2.064			
		df	*SS*	*MS*	*F*	*ignificance F*			**p-value**	0.197422	=F.DIST.RT(Q10, I35-I13, I13)		
	Regression	4	561.6047	140.4012	38.82785	7.24E-05							
	Residual	7	25.31194	3.615992									
	Total	11	586.9167										
		Coefficients	*andard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*				
	Intercept	14.34009	7.565093	1.89556	0.099856	-3.54851	32.22869	-3.54851	32.22869351				
	x1	2.290687	0.59909	3.82361	0.006512	0.874064	3.70731	0.874064	3.707309921				
	x2	0.516173	0.567794	0.909086	0.393528	-0.82645	1.858791	-0.82645	1.858791347				
	x3	-0.29052	0.153701	-1.89018	0.100649	-0.65397	0.072923	-0.65397	0.072923129				
	x4	0.006391	0.189974	0.033642	0.974102	-0.44283	0.455609	-0.44283	0.45560912				
	SUMMARY OUTPUT												
	Regression Statistics												
	Multiple R	0.965112											
	R Square	0.931441											
	Adjusted R	0.916206											
	Standard E	2.114459											
	Observatio	12											
Reduced Model	ANOVA												
		df	*SS*	*MS*	*F*	*ignificance F*							
	Regression	2	546.6782	273.3391	61.13686	5.78E-06							
	Residual	9	40.23844	4.470938									
	Total	11	586.9167										

The p-value turns out to be 0.1974.

Since this p-value is not less than .05, we will fail to reject the null hypothesis. This means we don't have sufficient evidence to say that either of the predictor variables *x3* or *x4* are statistically significant.

In other words, adding *x3* and *x4* to the regression model do not significantly improve the fit of the model.