How can I do simple main effects using anovalator in Stata?

Authored by

stats writer

July 1, 2024

Anovalator is a statistical software program used in Stata for conducting analysis of variance (ANOVA). It allows users to explore the relationships between multiple variables and identify significant differences between groups. One possible use of Anovalator is to conduct simple main effects analysis, which involves examining the effect of one independent variable on the dependent variable while holding other independent variables constant. This can be done by specifying the main effects of the variables in the model and using the built-in Anovalator command in Stata. By following the appropriate syntax and interpreting the results, users can easily perform simple main effects analysis and gain valuable insights into the relationships between variables in their data set.

How can I do simple main effects using anovalator? Stata FAQ

This page will demonstrate the use of the anovalator command (search anovalator) to compute simple main effects for both a two-factor model and a three-factor model. We will be using anovalator with the anova command on this page but anovalator works equally well with regress, xtmixed and many other estimation commands.

Example 1: Two-factor model

For our first example we will use the hsbanova dataset.

use https://stats.idre.ucla.edu/stat/data/hsbanova, clear

anova write grp##female

Number of obs = 200 R-squared = 0.2872

Root MSE = 8.14699 Adj R-squared = 0.2612

Source | Partial SS df MS F Prob > F

-----------+--

Model | 5135.17494 7 733.59642 11.05 0.0000

|

grp | 3641.68311 3 1213.89437 18.29 0.0000

female | 984.377328 1 984.377328 14.83 0.0002

grp#female | 575.513416 3 191.837805 2.89 0.0367

|

Residual | 12743.7001 192 66.3734378

-----------+--

Total | 17878.875 199 89.843593

You will note that the grp by female interaction is statistically significant.
We will show two ways of using anovalator to test the effect of grp for each
level of female. The first method uses the simple option in anovalator.
This approach will only work with two-factor models. We will also include the fratio
option in all of our examples since the disturbances are

assumed to be normally distributed in our anova model.

anovalator grp female, simple fratio

anovalator test of simple main effects for grp at(female=0)
chi2(3) = 46.001924 p-value = 5.666e-10
scaled as F-ratio = 15.333975

anovalator test of simple main effects for grp at(female=1)
chi2(3) = 13.644845 p-value = .00343069
scaled as F-ratio = 4.5482816

We can obtain the same results using anovalator with the maineffect
(abbreviated main) and at options.

anovalator grp, main fratio at(female=0)

anovalator main-effect for grp at(female=0)
chi2(3) = 46.001924 p-value = 5.666e-10
scaled as F-ratio = 15.333975

anovalator grp, main fratio at(female=1)

anovalator main-effect for grp at(female=1)
chi2(3) = 13.644845 p-value = .00343069
scaled as F-ratio = 4.5482816

Since tests of simple main effects are a type of post-hoc comparison we need to use an adjusted critical value to assess statistical significance. We will use the smecriticalvalue command (search smecriticalvalue) for this purpose. The smecriticalvalue command needs four pieces of information: n - the number of test performed; df1 - the degrees of freedom for each test; df2 - the degrees of freedon for the error term in the anova model; and dfmodel - the total degrees of freedom for all the terms in the model.

smecriticalvalue, n(2) df1(3) df2(192) dfmodel(7)

number of tests: 2
numerator df: 3
denominator df: 192
original model df: 7

Critical value of F for alpha = .05 using ...

Dunn's procedure = 3.3028802
Marascuilo & Levin = 3.6129331
per family error rate = 3.1847981
simultaneous test procedure = 3.6365283

By any of the four criteria above both tests of simple main effects are statistically significant.

Example 2: Three-factor model

Next we will run a three-factor model using the threeway dataset.

use https://stats.idre.ucla.edu/stat/data/threeway, clear

anova y a##b##c

Number of obs = 24 R-squared = 0.9689
Root MSE = 1.1547 Adj R-squared = 0.9403

Source | Partial SS df MS F Prob > F
-----------+--
Model | 497.833333 11 45.2575758 33.94 0.0000
|
a | 150 1 150 112.50 0.0000

b | .666666667 1 .666666667 0.50 0.4930

a#b | 160.166667 1 160.166667 120.13 0.0000

c | 127.583333 2 63.7916667 47.84 0.0000

a#c | 18.25 2 9.125 6.84 0.0104

b#c | 22.5833333 2 11.2916667 8.47 0.0051

a#b#c | 18.5833333 2 9.29166667 6.97 0.0098

|

Residual | 16 12 1.33333333

-----------+--

Total | 513.833333 23 22.3405797

The a#b#c threeway interaction is statistically significant. Prior research has suggested that we look at the b#c interaction for each level of a. Because this is a three-factor design we cannot use the simple option so we will use a variation of the second method shown above, this time using the twoway (abbreviated two) option.

anovalator grp female, two fratio at(a=1)

anovalator two-way interaction for b#c at(a=1)

chi2(2) = 30.5 p-value = 2.382e-07

scaled as F-ratio = 15.25

anovalator b c, two fratio at(a=2)

anovalator two-way interaction for b#c at(a=2)

chi2(2) = .375 p-value = .82902912

scaled as F-ratio = .1875

Since the b#c at a=1 is significant we will follow this analysis up

by looking at c for each level of b while keeping a at level one.

anovalator c, main fratio at(a=1 b=1)

anovalator main-effect for c at(a=1 b=1)

chi2(2) = 48 p-value = 3.775e-11

scaled as F-ratio = 24

anovalator c, main fratio at(a=1 b=2)

anovalator main-effect for c at(a=1 b=2)

chi2(2) = 1 p-value = .60653066

scaled as F-ratio = .5

We will use the smecriticalvalue command once again, this including all four tests since they have the same degrees of freedom.

smecriticalvalue, n(4) df1(2) df2(12) dfmodel(11)

number of tests: 4
numerator df: 2
denominator df: 12
original model df: 11

Critical value of F for alpha = .05 using ...

Dunn's procedure = 6.2753765
Marascuilo & Levin = 7.1335873
per family error rate = 6.4546898
simultaneous test procedure = 10.245969

There really wasn't much mystery here. Clearly, the F-ratios of .1875 and .5 cannot be significant.
While the F-ratios of 15.25 and 24 both exceed the most stringent criteria, the simultaneous test procedure.