How can I do ANOVA contrasts in Stata?

Authored by

stats writer

July 1, 2024

ANOVA contrasts in Stata refer to the statistical method used to compare the means of three or more groups. This method allows researchers to determine if there are significant differences between the groups and which specific group(s) differ from the others. To conduct ANOVA contrasts in Stata, researchers must first input their data and specify the desired contrast using the "contrast" command. Stata then calculates the contrast coefficients and provides the results in a table, including the estimated difference between groups and its corresponding p-value. This allows for a more comprehensive analysis of group differences and provides valuable insights for further statistical analysis. Overall, ANOVA contrasts in Stata provide a powerful tool for researchers to understand and compare group differences in their data.

How can I do ANOVA contrasts in Stata? | Stata FAQ

Stata does not have a built-in contrast command; however, ATS has developed a program that will do ANOVA contrasts. You can download the program anovacontrast.ado by typing search anovacontrast (see How can I use the search command to search for programs and get additional help? for more information about using search).

Now, let's read in an example dataset, crf24, adapted from Kirk (1968, First Edition).

use https://stats.idre.ucla.edu/stat/stata/faq/crf24

These data are from a 2×4 factorial design but the same

**data can also be used for
one-way ANOVA examples. The variable y is the dependent variable. The variable
a is an
independent variable with two levels while b is an independent variable with four levels.**

Using the anovacontrast command in a one-way ANOVA

```
anova y b

Number of obs = 32 R-squared = 0.8259
Root MSE = 1.21008 Adj R-squared = 0.8072

Source | Partial SS df MS F Prob > F
-----------+----------------------------------------------------
Model | 194.50 3 64.8333333 44.28 0.0000
|
b | 194.50 3 64.8333333 44.28 0.0000
|
Residual | 41.00 28 1.46428571
-----------+----------------------------------------------------
Total | 235.50 31 7.59677419

table b, contents(mean y)
```

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

b | mean(y)

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

1 | 2.75

2 | 3.5

3 | 6.25

4 | 9

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

It is quite clear that there is a significant overall F for the
independent variable b. Now, let's devise some contrasts that we can test:
1) group 3 versus group 4
2) the average of groups 1 and 2 versus the average of groups 3 and 4
3) the average of groups 1, 2, and 3 versus group 4

anovacontrast b, values(0 0 1 -1)

Contrast variable b (0 0 1 -1) Dep Var = y
source SS df MS N of obs = 32
---------+---------------------------------- F = 20.66
contrast | 30.25 1 30.2500 Prob > F = 0.0001

error | 41 28 1.4643
---------+----------

anovacontrast b, values(1 1 -1 -1)

Contrast variable b (1 1 -1 -1) Dep Var = y
source SS df MS N of obs = 32
---------+---------- F = 110.63
contrast | 162 1 162.0000 Prob > F = 0.0000
error | 41 28 1.4643
---------+----------

anovacontrast b, values(1 1 1 -3)

Contrast variable b (1 1 1 -3) Dep Var = y
source SS df MS N of obs = 32
---------+---------- F = 95.72
contrast | 140.166667 1 140.1667 Prob > F = 0.0000
error | 41 28 1.4643
---------+----------

Using the anovacontrast command in a two-way ANOVA

Now let's try the same contrasts on b but in a two-way ANOVA.

```
anova y a b a*b

Number of obs = 32 R-squared = 0.9214
Root MSE = .877971 Adj R-squared = 0.8985

Source | Partial SS df MS F Prob > F
-----------+----------------------------------------------------
Model | 217.00 7 31.00 40.22 0.0000
|
a | 3.125 1 3.125 4.05 0.0554
b | 194.50 3 64.8333333 84.11 0.0000
a*b | 19.375 3 6.45833333 8.38 0.0006
|
Residual | 18.50 24 .770833333
-----------+----------------------------------------------------
Total | 235.50 31 7.59677419

anovacontrast b, values(0 0 1 -1)

Contrast variable b (0 0 1 -1) Dep Var = y
source SS df MS N of obs = 32
---------+------------------------------ F = 39.24
contrast | 30.25 1 30.2500 Prob > F = 0.0000
error | 18.5 24 0.7708
---------+------------------------------
```

```
anovacontrast b, values(1 1 -1 -1)

Contrast variable b (1 1 -1 -1) Dep Var = y
source SS df MS N of obs = 32
---------+------------------------------ F = 210.16
contrast | 162 1 162.0000 Prob > F = 0.0000
error | 18.5 24 0.7708
---------+------------------------------

anovacontrast b, values(1 1 1 -3)

Contrast variable b (1 1 1 -3) Dep Var = y
source SS df MS N of obs = 32
---------+------------------------------ F = 181.84
contrast | 140.166667 1 140.1667 Prob > F = 0.0000
error | 18.5 24 0.7708
---------+------------------------------
```

Note that the F-ratios in these contrasts are larger than the F-ratios in the one-way ANOVA example. This is because the two-way ANOVA has a smaller mean square residual than the one-way ANOVA.