oneway — One-way analysis of variance

Description | Quick start | Menu | Syntax
Options | Remarks and examples | Stored results | Methods and formulas
References | Also see

Description

The `oneway` command reports one-way analysis-of-variance (ANOVA) models and performs multiple-comparison tests.

If you wish to fit more complicated ANOVA layouts or wish to fit analysis-of-covariance (ANCOVA) models, see [R] **anova**.

See [D] **encode** for examples of fitting ANOVA models on string variables.

See [R] **loneway** for an alternative `oneway` command with slightly different features.

Quick start

One-way ANOVA model of `y` for factor `a`

```
oneway y a
```

Report the mean and std. dev. of `y` and number of observations for each level of `a`

```
oneway y a, tabulate
```

Report all pairwise comparisons of the means of `y` across levels of `a` with p-values adjusted using Bonferroni's procedure

```
oneway y a, bonferroni
```

Same as above, but adjust p-values for multiple comparisons using Scheffé's method

```
oneway y a, scheffe
```

Menu

Statistics > Linear models and related > ANOVA/MANOVA > One-way ANOVA

Syntax

oneway *response_var factor_var* [*if*] [*in*] [*weight*] [, *options*]

options	Description
Main	
bonferroni	Bonferroni multiple-comparison test
scheffe	Scheffé multiple-comparison test
sidak	Šidák multiple-comparison test
tabulate	produce summary table
[no]means	include or suppress means; default is means
[no]standard	include or suppress standard deviations; default is standard
[no]freq	include or suppress frequencies; default is freq
[no]obs	include or suppress number of obs; default is obs if data are weighted
noanova	suppress the ANOVA table
nolabel	show numeric codes, not labels
wrap	do not break wide tables
missing	treat missing values as categories

by and collect are allowed; see **[U] 11.1.10 Prefix commands**.

aweights and fweights are allowed; see **[U] 11.1.6 weight**.

Options

Main

bonferroni reports the results of a Bonferroni multiple-comparison test.

scheffe reports the results of a Scheffé multiple-comparison test.

sidak reports the results of a Šidák multiple-comparison test.

tabulate produces a table of summary statistics of the *response_var* by levels of the *factor_var*. The table includes the mean, standard deviation, frequency, and, if the data are weighted, the number of observations. Individual elements of the table may be included or suppressed by using the [no]means, [no]standard, [no]freq, and [no]obs options. For example, typing

```
oneway response factor, tabulate means standard
```

produces a summary table that contains only the means and standard deviations. You could achieve the same result by typing

```
oneway response factor, tabulate nofreq
```

[no]means includes or suppresses only the means from the table produced by the tabulate option. See tabulate above.

[no]standard includes or suppresses only the standard deviations from the table produced by the tabulate option. See tabulate above.

[no]freq includes or suppresses only the frequencies from the table produced by the tabulate option. See tabulate above.

[no]obs includes or suppresses only the reported number of observations from the table produced by the tabulate option. If the data are not weighted, only the frequency is reported. If the data are weighted, the frequency refers to the sum of the weights. See tabulate above.

noanova suppresses the display of the ANOVA table.

nolabel causes the numeric codes to be displayed rather than the value labels in the ANOVA and multiple-comparison test tables.

wrap requests that Stata not break up wide tables to make them more readable.

missing requests that missing values of *factor_var* be treated as a category rather than as observations to be omitted from the analysis.

Remarks and examples

Remarks are presented under the following headings:

Introduction
Obtaining observed means
Multiple-comparison tests
Weighted data
Video example

Introduction

The oneway command reports one-way ANOVA models. To perform a one-way layout of a variable called endog on exog, type oneway endog exog.

▷ Example 1

We run an experiment varying the amount of fertilizer used in growing apple trees. We test four concentrations, using each concentration in three groves of 12 trees each. Later in the year, we measure the average weight of the fruit.

If all had gone well, we would have had 3 observations on the average weight for each of the four concentrations. Instead, two of the groves were mistakenly leveled by a confused man on a large bulldozer. We are left with the following dataset:

```
. use https://www.stata-press.com/data/r19/apple
(Apple trees)
. describe
Contains data from https://www.stata-press.com/data/r19/apple.dta
 Observations:            10                  Apple trees
    Variables:             2                  16 Jan 2024 11:23
-------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------
treatment       int     %8.0g                 Fertilizer
weight          double  %10.0g                Average weight in grams
-------------------------------------------------------------------------------
Sorted by:
```

```
. list, abbreviate(10)
```

	treatment	weight
1.	1	117.5
2.	1	113.8
3.	1	104.4
4.	2	48.9
5.	2	50.4
6.	2	58.9
7.	3	70.4
8.	3	86.9
9.	4	87.7
10.	4	67.3

To obtain the one-way ANOVA results, we type

```
. oneway weight treatment
                        Analysis of variance
    Source              SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between groups      5295.54433      3   1765.18144     21.46     0.0013
 Within groups      493.591667      6   82.2652778
------------------------------------------------------------------------
    Total             5789.136      9   643.237333
Bartlett's equal-variances test: chi2(3) =   1.3900    Prob>chi2 = 0.708
```

We find significant (at better than the 1% level) differences among the four concentrations.

◁

❑ Technical note

Rather than using the `oneway` command, we could have performed this analysis by using `anova`. Example 1 in [R] **anova** repeats this same analysis. You may wish to compare the output.

You will find the `oneway` command quicker than the `anova` command, and, as you will learn, `oneway` allows you to perform multiple-comparison tests. On the other hand, `anova` will let you generate predictions, examine the covariance matrix of the estimators, and perform more general hypothesis tests.

❑

❑ Technical note

Although the output is a usual ANOVA table, let's run through it anyway. The between-group sum of squares for the model is 5295.5 with 3 degrees of freedom, resulting in a mean square of $5295.5/3 \approx 1765.2$. The corresponding F statistic is 21.46 and has a significance level of 0.0013. Thus, the model appears to be significant at the 0.13% level.

The second line summarizes the within-group (residual) variation. The within-group sum of squares is 493.59 with 6 degrees of freedom, resulting in a mean squared error of 82.27.

The between- and residual-group variations sum to the total sum of squares (TSS), which is reported as 5789.1 in the last line of the table. This is the TSS of `weight` after removal of the mean. Similarly, the between plus residual degrees of freedom sum to the total degrees of freedom, 9. Remember that there are 10 observations. Subtracting 1 for the mean, we are left with 9 total degrees of freedom.

At the bottom of the table, Bartlett's test for equal variances is reported. The value of the statistic is 1.39. The corresponding significance level (χ^2 with 3 degrees of freedom) is 0.708, so we cannot reject the assumption that the variances are homogeneous.

❑

Obtaining observed means

▷ Example 2

We typed `oneway weight treatment` to obtain an ANOVA table of weight of fruit by fertilizer concentration. Although we obtained the table, we obtained no information on which fertilizer seems to work the best. If we add the `tabulate` option, we obtain that additional information:

```
. oneway weight treatment, tabulate
```

Fertilizer	Summary of Average weight in grams Mean	Std. dev.	Freq.
1	111.9	6.7535176	3
2	52.733333	5.3928966	3
3	78.65	11.667262	2
4	77.5	14.424978	2
Total	80.62	25.362124	10

Analysis of variance

Source	SS	df	MS	F	Prob > F
Between groups	5295.54433	3	1765.18144	21.46	0.0013
Within groups	493.591667	6	82.2652778		
Total	5789.136	9	643.237333		

Bartlett's equal-variances test: chi2(3) = 1.3900 Prob>chi2 = 0.708

We find that the average weight was largest when we used fertilizer concentration 1.

◁

Multiple-comparison tests

▷ Example 3: Bonferroni multiple-comparison test

oneway can also perform multiple-comparison tests using either Bonferroni, Scheffé, or Šidák normalizations. For instance, to obtain the Bonferroni multiple-comparison test, we specify the bonferroni option:

```
. oneway weight treatment, bonferroni
                        Analysis of variance
    Source              SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between groups      5295.54433      3   1765.18144     21.46     0.0013
 Within groups      493.591667      6   82.2652778
------------------------------------------------------------------------
    Total             5789.136      9   643.237333
Bartlett's equal-variances test: chi2(3) =   1.3900    Prob>chi2 = 0.708
             Comparison of Average weight in grams by Fertilizer
                                (Bonferroni)
Row Mean-
Col Mean |          1          2          3
---------+---------------------------------
       2 |   -59.1667
         |      0.001
         |
       3 |     -33.25    25.9167
         |      0.042      0.122
         |
       4 |      -34.4    24.7667      -1.15
         |      0.036      0.146      1.000
```

The results of the Bonferroni test are presented as a matrix. The first entry, −59.17, represents the difference between fertilizer concentrations 2 and 1 (labeled "Row Mean - Col Mean" in the upper stub of the table). Remember that in the previous example we requested the tabulate option. Looking back, we find that the means of concentrations 1 and 2 are 111.90 and 52.73, respectively. Thus, 52.73 − 111.90 = −59.17.

Underneath that number is reported "0.001". This is the Bonferroni-adjusted significance of the difference. The difference is significant at the 0.1% level. Looking down the column, we see that concentration 3 is also worse than concentration 1 (4.2% level), as is concentration 4 (3.6% level).

On the basis of this evidence, we would use concentration 1 if we grew apple trees.

◁

▷ Example 4: Scheffé multiple-comparison test

We can just as easily obtain the Scheffé-adjusted significance levels. Rather than specifying the `bonferroni` option, we specify the `scheffe` option.

We will also add the `noanova` option to prevent Stata from redisplaying the ANOVA table:

```
. oneway weight treatment, noanova scheffe

            Comparison of Average weight in grams by Fertilizer
                                    (Scheffe)
Row Mean-
Col Mean |          1          2          3
---------+------------------------------------
       2 |   -59.1667
         |      0.001
         |
       3 |     -33.25    25.9167
         |      0.039      0.101
         |
       4 |      -34.4    24.7667      -1.15
         |      0.034      0.118      0.999
```

The differences are the same as those we obtained in the Bonferroni output, but the significance levels are not. According to the Bonferroni-adjusted numbers, the significance of the difference between fertilizer concentrations 1 and 3 is 4.2%. The Scheffé-adjusted significance level is 3.9%.

We will leave it to you to decide which results are more accurate.

◁

▷ Example 5: Šidák multiple-comparison test

Let's conclude this example by obtaining the Šidák-adjusted multiple-comparison tests. We do this to illustrate Stata's capabilities to calculate these results, because searching across adjustment methods until you find the results you want is not a valid technique for obtaining significance levels.

```
. oneway weight treatment, noanova sidak

            Comparison of Average weight in grams by Fertilizer
                                    (Sidak)
Row Mean-
Col Mean |          1          2          3
---------+------------------------------------
       2 |   -59.1667
         |      0.001
         |
       3 |     -33.25    25.9167
         |      0.041      0.116
         |
       4 |      -34.4    24.7667      -1.15
         |      0.035      0.137      1.000
```

We find results that are similar to the Bonferroni-adjusted numbers.

◁

Henry Scheffé (1907–1977) was born in New York. He studied mathematics at the University of Wisconsin, gaining a doctorate with a dissertation on differential equations. He taught mathematics at Wisconsin, Oregon State University, and Reed College, but his interests changed to statistics and he joined Wilks at Princeton. After periods at Syracuse, UCLA, and Columbia, Scheffé settled in Berkeley from 1953. His research increasingly focused on linear models and particularly ANOVA, on which he produced a celebrated monograph. His death was the result of a bicycle accident.

Weighted data

▷ Example 6

`oneway` can work with both weighted and unweighted data. Let's assume that we wish to perform a one-way layout of the deathrate on the four census regions of the United States using state data. Our data contain three variables, `drate` (the deathrate), `region` (the region), and `pop` (the population of the state).

To fit the model, we type `oneway drate region [weight=pop]`, although we typically abbreviate `weight` as `w`. We will also add the `tabulate` option to demonstrate how the table of summary statistics differs for weighted data:

```
. use https://www.stata-press.com/data/r19/census8
(1980 Census data by state)
. oneway drate region [w=pop], tabulate
(analytic weights assumed)
```

Census region	Summary of Deathrate Mean	Std. dev.	Freq.	Obs
NE	97.15	5.82	49135283	9
N Cntrl	88.10	5.58	58865670	12
South	87.05	10.40	74734029	16
West	75.65	8.23	43172490	13
Total	87.34	10.43	225907472	50

Analysis of variance

Source	SS	df	MS	F	Prob > F
Between groups	2360.92281	3	786.974272	12.17	0.0000
Within groups	2974.09635	46	64.6542685		
Total	5335.01916	49	108.877942		

```
Bartlett's equal-variances test: chi2(3) =   5.4971    Prob>chi2 = 0.139
```

When the data are weighted, the summary table has four columns rather than three. The column labeled "`Freq.`" reports the sum of the weights. The overall frequency tells us that there are approximately 226 million people in the United States, or at least there were in 1980.

The ANOVA table is appropriately weighted. Also see **[U] 11.1.6 weight**.

◁

Video example

One-way ANOVA in Stata

Stored results

oneway stores the following in r():

Scalars

r(N)	number of observations	r(df_m)	between-group degrees of freedom
r(F)	F statistic	r(rss)	within-group sum of squares
r(df_r)	within-group degrees of freedom	r(chi2bart)	Bartlett's χ^2
r(mss)	between-group sum of squares	r(df_bart)	Bartlett's degrees of freedom

Matrices

r(ANOVA)	matrix containing the sum of squares with their degrees of freedom, mean squares, F statistic, and p-value

Methods and formulas

Methods and formulas are presented under the following headings:

One-way analysis of variance
Bartlett's test
Multiple-comparison tests

One-way analysis of variance

The model of one-way ANOVA is

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

for levels $i = 1, \ldots, k$ and observations $j = 1, \ldots, n_i$. Define $\overline{y}_i$ as the (weighted) mean of y_{ij} over j and $\overline{y}$ as the overall (weighted) mean of y_{ij}. Define w_{ij} as the weight associated with y_{ij}, which is 1 if the data are unweighted. w_{ij} is normalized to sum to $n = \sum_i n_i$ if aweights are used and is otherwise not normalized. w_i refers to $\sum_j w_{ij}$, and w refers to $\sum_i w_i$.

The between-group sum of squares is then

$$S_1 = \sum_i w_i(\overline{y}_i - \overline{y})^2$$

The TSS is

$$S = \sum_i \sum_j w_{ij}(y_{ij} - \overline{y})^2$$

The within-group sum of squares is given by $S_e = S - S_1$.

The between-group mean square is $s_1^2 = S_1/(k-1)$, and the within-group mean square is $s_e^2 = S_e/(w-k)$. The test statistic is $F = s_1^2/s_e^2$. See, for instance, Snedecor and Cochran (1989).

Bartlett's test

Bartlett's test assumes that you have m independent, normal, random samples and tests the hypothesis $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_m^2$. The test statistic, M, is defined as

$$M = \frac{(T-m)\ln\hat{\sigma}^2 - \sum(T_i-1)\ln\hat{\sigma}_i^2}{1 + \frac{1}{3(m-1)}\left\{\left(\sum\frac{1}{T_i-1}\right) - \frac{1}{T-m}\right\}}$$

where there are T overall observations, T_i observations in the ith group, and

$$(T_i-1)\hat{\sigma}_i^2 = \sum_{j=1}^{T_i}(y_{ij}-\overline{y}_i)^2$$

$$(T-m)\hat{\sigma}^2 = \sum_{i=1}^{m}(T_i-1)\hat{\sigma}_i^2$$

An approximate test of the homogeneity of variance is based on the statistic M with critical values obtained from the χ^2 distribution of $m-1$ degrees of freedom. See Bartlett (1937) or Draper and Smith (1998, 56–57).

Multiple-comparison tests

Let's begin by reviewing the logic behind these adjustments. The "standard" t statistic for the comparison of two means is

$$t = \frac{\overline{y}_i - \overline{y}_j}{s\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}}$$

where s is the overall standard deviation, $\overline{y}_i$ is the measured average of y in group i, and n_i is the number of observations in the group. We perform hypothesis tests by calculating this t statistic. We simultaneously choose a critical level, α, and look up the t statistic corresponding to that level in a table. We reject the hypothesis if our calculated t exceeds the value we looked up. Alternatively, because we have a computer at our disposal, we calculate the significance level e corresponding to our calculated t statistic, and if $e < \alpha$, we reject the hypothesis.

This logic works well when we are performing one test. Now consider what happens when we perform several separate tests, say, n of them. Let's assume, just for discussion, that we set α equal to 0.05 and that we will perform six tests. For each test, we have a 0.05 probability of falsely rejecting the equality-of-means hypothesis. Overall, then, our chances of falsely rejecting *at least one* of the hypotheses is $1-(1-0.05)^6 \approx 0.26$ if the tests are independent.

The idea behind multiple-comparison tests is to control for the fact that we will perform multiple tests and to reduce our overall chances of falsely rejecting each hypothesis to α rather than letting our chances increase with each additional test. (See Miller [1981] and Hochberg and Tamhane [1987] for rather advanced texts on multiple-comparison procedures.)

The Bonferroni adjustment (see Miller [1981]; also see van Belle et al. [2004, 534–537]) does this by (falsely but approximately) asserting that the critical level we should use, a, is the true critical level, α, divided by the number of tests, n; that is, $a = \alpha/n$. For instance, if we are going to perform six tests, each at the 0.05 significance level, we want to adopt a critical level of $0.05/6 \approx 0.00833$.

We can just as easily apply this logic to e, the significance level associated with our t statistic, as to our critical level α. If a comparison has a calculated significance of e, then its "real" significance, adjusted for the fact of n comparisons, is $n \times e$. If a comparison has a significance level of, say, 0.012, and we perform six tests, then its "real" significance is 0.072. If we adopt a critical level of 0.05, we cannot reject the hypothesis. If we adopt a critical level of 0.10, we can reject it.

Of course, this calculation can go above 1, but that just means that there is no $\alpha < 1$ for which we could reject the hypothesis. (This situation arises because of the crude nature of the Bonferroni adjustment.) Stata handles this case by simply calling the significance level 1. Thus, the formula for the Bonferroni significance level is

$$e_b = \min(1, en)$$

where $n = k(k-1)/2$ is the number of comparisons.

The Šidák adjustment (Šidák [1967]; also see Winer, Brown, and Michels [1991, 165–166]) is slightly different and provides a tighter bound. It starts with the assertion that

$$a = 1 - (1-\alpha)^{1/n}$$

Turning this formula around and substituting calculated significance levels, we obtain

$$e_s = \min\left\{1, 1-(1-e)^n\right\}$$

For example, if the calculated significance is 0.012 and we perform six tests, the "real" significance is approximately 0.07.

The Scheffé test (Scheffé [1953, 1959]; also see Kuehl [2000, 97–98]) differs in derivation, but it attacks the same problem. Let there be k means for which we want to make all the pairwise tests. Two means are declared significantly different if

$$t \geq \sqrt{(k-1)F(\alpha; k-1, \nu)}$$

where $F(\alpha; k-1, \nu)$ is the α-critical value of the F distribution with $k-1$ numerator and ν denominator degrees of freedom. Scheffé's test has the nicety that it never declares a contrast significant if the overall F test is not significant.

Turning the test around, Stata calculates a significance level

$$\hat{e} = F\left(\frac{t^2}{k-1}, k-1, \nu\right)$$

For instance, you have a calculated t statistic of 4.0 with 50 degrees of freedom. The simple t test says that the significance level is 0.00021. The F test equivalent, 16 with 1 and 50 degrees of freedom, says the same. If you are comparing three means, however, you calculate an F test of 8.0 with 2 and 50 degrees of freedom, which says that the significance level is 0.0010.

References

Acock, A. C. 2023. *A Gentle Introduction to Stata*. Rev. 6th ed. College Station, TX: Stata Press.

Altman, D. G. 1991. *Practical Statistics for Medical Research*. London: Chapman and Hall/CRC.

Bartlett, M. S. 1937. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society*, A ser., 160: 268–282.

Daniel, C., and E. L. Lehmann. 1979. Henry Scheffé 1907–1977. *Annals of Statistics* 7: 1149–1161. https://doi.org/10.1214/aos/1176344837.

Draper, N., and H. Smith. 1998. *Applied Regression Analysis*. 3rd ed. New York: Wiley. https://doi.org/10.1002/9781118625590.

Hochberg, Y., and A. C. Tamhane. 1987. *Multiple Comparison Procedures*. New York: Wiley.

Kuehl, R. O. 2000. *Design of Experiments: Statistical Principles of Research Design and Analysis*. 2nd ed. Belmont, CA: Duxbury.

Marchenko, Y. V. 2006. Estimating variance components in Stata. *Stata Journal* 6: 1–21.

Miller, R. G., Jr. 1981. *Simultaneous Statistical Inference*. 2nd ed. New York: Springer. https://doi.org/10.1007/978-1-4613-8122-8.

Scheffé, H. 1953. A method for judging all contrasts in the analysis of variance. *Biometrika* 40: 87–104. https://doi.org/10.2307/2333100.

———. 1959. *The Analysis of Variance*. New York: Wiley.

Šidák, Z. 1967. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association* 62: 626–633. https://doi.org/10.2307/2283989.

Snedecor, G. W., and W. G. Cochran. 1989. *Statistical Methods*. 8th ed. Ames, IA: Iowa State University Press.

van Belle, G., L. D. Fisher, P. J. Heagerty, and T. S. Lumley. 2004. *Biostatistics: A Methodology for the Health Sciences*. 2nd ed. New York: Wiley.

Winer, B. J., D. R. Brown, and K. M. Michels. 1991. *Statistical Principles in Experimental Design*. 3rd ed. New York: McGraw–Hill.

Also see

[R] **anova** — Analysis of variance and covariance

[R] **loneway** — Large one-way ANOVA, random effects, and reliability

[PSS-2] **power oneway** — Power analysis for one-way analysis of variance

Stata, Stata Press, Mata, NetCourse, and NetCourseNow are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow is a trademark of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.

For suggested citations, see the FAQ on citing Stata documentation.